

# Validation of Patient-Reported Outcomes for Clinical Trials in Allergic Rhinitis: A Systematic Review



Moises A. Calderón, MD, PhD<sup>a</sup>, Thomas B. Casale, MD<sup>b</sup>, and Pascal Demoly, MD, PhD<sup>c,d</sup> *London, United Kingdom; Tampa, Fla; and Montpellier and Paris, France*

Although regulatory authorities have recently recommended the use of a combined symptom-medication score as a primary efficacy end point, none has been psychometrically validated. Here, we sought to determine to what extent allergic rhinitis (AR)-related patient-reported outcomes (symptom scores, medication scores, disease control scores, and satisfaction or quality-of-life scales) have been assessed for construct, content, and/or criterion validity, reliability, responsiveness, and the minimal clinically important difference. We searched the PubMed database from January 1997 to June 2018 with logical combinations of key words related to validation, AR, and patient-rated outcomes and scales. From a total of 1705 potentially relevant publications, 55 were reviewed. Despite the current emphasis on a combined symptom-medication score for evaluating the efficacy of allergen immunotherapy in AR, symptom scores have not been extensively validated, and we did not find any publications describing the validation of a medication score. Disease control scales (mainly the Rhinitis Control Assessment Test, the Control of Allergic Rhinitis and Asthma Test, and the Allergic Rhinitis Control Test) and health-related quality-of-life scales (mainly the Rhinoconjunctivitis Quality of Life Questionnaire [RQLQ] and the mini-RQLQ)

have been extensively validated in AR but have some practical disadvantages as primary efficacy criteria in clinical trials. © 2019 Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology (J Allergy Clin Immunol Pract 2019;7:1450-61)

**Key words:** Allergen immunotherapy; Allergic rhinitis; Score; Validation; Psychometric; Reliability; Responsiveness; Minimally clinical important difference

## INTRODUCTION

### The need for valid patient-reported outcomes

Allergic rhinoconjunctivitis (AR) can impair respiratory health, quality of life (QOL), health care costs, social activities, sleep, and academic/work performance<sup>1-4</sup> and can even increase the risk of traffic accidents.<sup>5</sup> Furthermore, AR is acknowledged to be a step in the “allergic march” toward allergic asthma, with its potentially life-threatening and enormous societal cost consequences.<sup>6,7</sup> Although pharmacologic (symptomatic) medications can provide short-term symptom relief,<sup>8</sup> they may not be effective or well tolerated by some patients—especially those with moderate to severe AR—and they do not treat the underlying disease. The only currently available disease-modifying treatment for both AR and allergic asthma (ie, with the potential to increase allergen tolerance and at least slow the allergic march) is allergen immunotherapy (AIT).<sup>9-17</sup>

In major markets such as the United States and the European Union, developers of new AIT formulations and other treatments for atopic diseases must demonstrate safety and efficacy in pivotal phase III double-blind, placebo-controlled, randomized clinical trials (DBPC RCTs).<sup>18</sup> Safety and efficacy outcomes are influenced by many key aspects of clinical trial design, implementation, and analysis. In the field of atopic disease, it is clear that many of these aspects require further optimization.<sup>10,19-27</sup> The choice of the criterion used to evaluate efficacy is crucial.<sup>10,25,28,29</sup> From an ethical standpoint, one can even consider (somewhat provocatively) that the use of a nonvalidated patient-reported outcome (PRO) criterion to evaluate efficacy in a group of patients could be equated to the use of a nonvalidated (ie, nonapproved) imaging method, functional testing method, or medical device in clinical practice. Hence, establishing the validity of a PRO is ethically, legally, and clinically important.

In the absence of a “gold standard” criterion, a large number of patient-reported scales, scores, and instruments have been used to measure efficacy in the treatment of AR with AIT and/or symptomatic medications: nasal symptom scores, ocular symptom scores, total (nasal + ocular) symptom scores, individual

<sup>a</sup>National Heart & Lung Institute, Imperial College London, Royal Brompton Hospital NHS, London, United Kingdom

<sup>b</sup>Division of Allergy and Immunology, University of South Florida, Tampa, Fla

<sup>c</sup>Division of Allergy, Department of Pulmonology, Hôpital Anaud de Villeneuve, University Hospital of Montpellier, Montpellier, France

<sup>d</sup>Sorbonne Universités, UMR-S 1136, IPLESP, Equipe EPAR, Paris, France

This study was funded jointly by ALK-Abelló A/S (Hørsholm, Denmark) and Stallergenes Greer (London, UK) for copyediting services by David Fraser (Biotech Communication SARL, Ploudalmezeau, France).

Conflicts of interest: M. A. Calderón has received lecture fees from ALK, Stallergenes Greer, Merck, Allergopharma, and HAL Allergy; and has received consultancy fees from ALK, Stallergenes Greer, Merck, HAL Allergy, and ASIT biotech. T. B. Casale has been an investigator and consultant for Novartis and Genentech with all fees to his university employer. P. Demoly has received consultancy fees from ALK, Stallergenes Greer, Allergy Therapeutics, YSlab, Sanofi, Bausch & Lomb, AstraZeneca, and ThermoFisher Scientific; has received fees for participation in review activities from BTT; and has received lecture fees from MYLAN, Chiesi, ALK, and Stallergenes Greer.

Received for publication October 27, 2018; revised January 5, 2019; accepted for publication January 7, 2019.

Available online February 20, 2019.

Corresponding author: Pascal Demoly, MD, PhD, Division of Allergy, Department of Pulmonology, Hôpital Anaud de Villeneuve, University Hospital of Montpellier, 371, av du Doyen G. Giraud, Montpellier CEDEX 5 34295, France. E-mail: [pascal.demoly@inserm.fr](mailto:pascal.demoly@inserm.fr).

2213-2198

© 2019 Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology

<https://doi.org/10.1016/j.jaip.2019.01.015>

#### Abbreviations used

ACS- Allergy-Control-SCORE  
AIT- Allergen immunotherapy  
AOS- Allergy Outcome Survey  
AR- Allergic rhinitis  
ARCT- Allergic Rhinitis Control Test  
CARAT10- Control of Allergic Rhinitis and Asthma Test  
COSMIN- Consensus-Based Standards for the Selection of Health Measurement Instruments  
CQ5- Congestion Quantifier 5-Item Screener  
CQ7- Congestion Quantifier 7-Item Test  
CSMS- Combined symptom-medication score  
DBPC RCT- Double-blind, placebo-controlled randomized clinical trial  
EMA- European Medicines Agency  
MCID- Minimal clinically important difference  
NOSE- Nasal Obstruction Symptom Evaluation  
PRO- Patient-reported outcome  
QOL- Quality of life  
RCAT- Rhinitis Control Assessment Test  
RCSS- Rhinitis Control Scoring System  
RQLQ- Rhinoconjunctivitis Quality of Life Questionnaire  
RTSS- Rhinoconjunctivitis Total Symptom Score  
T5SS- 5-item symptom score  
T6SS- 6-item symptom score  
VAS- visual analogue scale

symptom scores, medication-adjusted symptom scores, medication scores, QOL scores, combined symptom-medication scores (CSMSs), visual analogue scales (VASs) of overall symptom burden, disease control questionnaires, physician- and patient-rated improvement/worsening, symptom-free days, “well days,” “worst days,” satisfaction questionnaires, nasal provocation tests, and conjunctival provocation tests (for comprehensive reviews, see studies by Nelson et al<sup>25</sup> and Pfaar et al<sup>28,29</sup>).

Regulatory authorities in both the United States and Europe have expressed a preference for a CSMS that reflects both symptom severity and rescue medication intake.<sup>30,31</sup> However, a European Medicines Authority (EMA) guideline developed over the period 2007 to 2008 pointed out that “Up to now, no validated symptom score exists, but the measurement of symptoms on a 4-point rating scale ... is generally accepted .... Likewise, no validated medication score exists .... So far, no validated system for balancing symptom and medication score exist.”<sup>30</sup>

Outcomes for AIT trials have been described and commented on in depth.<sup>25,28,29</sup> However, we decided to focus on the 3 main quality domains applicable to PROs: validity, reliability, and responsiveness (Figure 1). These are described in more detail herein. In the present review, we followed the definitions of validity (subdivided into content validity, construct validity, and criterion validity) given in the recent Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN).<sup>32-35</sup>

#### Content validity

The COSMIN definition of *content validity* is “the degree to which the content of an instrument is an adequate reflection of the construct to be measured” (ie, the comprehensiveness of the PRO).<sup>33</sup> In the present review, we considered that a group of researchers had assessed content validity if they had (1) explicitly

stated this or (2) made a systematic or semisystematic attempt to select and then test an instrument’s constituent items.

#### Construct validity

According to COSMIN, *construct validity* is “the degree to which the scores of an instrument are consistent with hypotheses (e.g. with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups).”<sup>33</sup> It notably “includes structural validity, hypothesis testing, and cross-cultural validity.” In the present review, we considered that a group of researchers had assessed construct validity if they had (1) explicitly stated this, (2) related the AR score under study to other AR scores, or (3) looked at whether the AR score under study differentiated between a group of patients with AR (as defined by a physician, for example) and healthy controls.

#### Criterion validity

A PRO is said to have *criterion validity* if its results are an adequate reflection of a criterion standard.<sup>33</sup> However, there is debatably no universally accepted criterion standard PRO for AR symptom and disease control scores, and there is overlap between criterion validity and construct validity. In the present review, we considered that a group of researchers had assessed criterion validity if they had explicitly stated this.

#### Reliability and responsiveness

*Reliability* is defined by COSMIN as “the extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health-related PRO (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater).”<sup>33</sup> In the present review, we considered that a group of researchers had assessed reliability if they had performed test-retest, interrater, and intrarater analyses or studied internal consistency (Figure 1).<sup>33</sup> *Responsiveness* is the ability of a health-related PRO instrument to detect a change over time in the construct to be measured.<sup>33</sup> A key component of responsiveness is the minimal clinically important difference (MCID), which denotes the smallest score or change in score likely to be important from the patient’s perspective.<sup>36-38</sup> The MCID can be determined using an anchor-based method (often relative to a 15-point global rating of change scale or the change in the Rhinoconjunctivitis Quality of Life Questionnaire [RQLQ]), a distribution-based method (often 0.5 of the SD of the mean PRO value), or an expert consensus.<sup>36-40</sup> We considered that a group of researchers had assessed responsiveness if they had assessed the MCID or, more generally, the ability of an instrument to detect a change over time in the construct to be measured.

Hence, the objective of the present review was to contribute to the ongoing debate on validated efficacy criteria in AIT trials by determining to what extent AR-related PROs (symptom scales, medication scales, disease control scales, and rhinitis-specific QOL scales) have been validated, that is, assessed for validity, reliability, and responsiveness.<sup>41,42</sup>

#### METHODS

Using the US National Library of Medicine’s online tool, we searched the PubMed database from January 1997 to June 2018 with logical combinations of the following key words: allerg\*, rhin\*, valid\*, validat\*, score\*, scale\*, control\*, outcome\*, psychometr\*,

|                                       |                |                                                                                                        |
|---------------------------------------|----------------|--------------------------------------------------------------------------------------------------------|
| Quality of a patient-reported outcome | Validity       | Construct validity<br>(convergent or discriminant, cross-sectional or longitudinal, language)          |
|                                       |                | Content validity<br>(comprehensiveness of item selection)                                              |
|                                       |                | Criterion validity<br>(vs. a gold standard or an outcome)                                              |
|                                       | Reliability    | Internal consistency                                                                                   |
|                                       |                | Test-retest, inter-rater and intra-rater reliability                                                   |
|                                       |                | Measurement error<br>(systematic and random errors in a score that are not attributed to true changes) |
|                                       | Responsiveness | Responsiveness to change<br>(including the minimal clinically important difference)                    |

**FIGURE 1.** The main measurement properties of instruments for the evaluation of health-related PROs.

MCID, MID, relevan\*, anchor-based, distribution-based, VAS, quality, and QOL. The search was limited to publications in English. Review articles, position papers, “rostrum” articles, and abstract publications were also excluded. The search period was limited to the last 20 years because this corresponds to the period for the clinical development and marketing of “modern” AIT formulations, such as sublingual tablets. Abstracts were obtained and screened for relevance. If an abstract was found to be relevant, the full publication was retrieved and read. The reference lists of retrieved full publications were also screened for other relevant publications. Finally, we excluded publications on instruments that could not, in our opinion, be scored on a daily or even weekly basis before and then throughout a pivotal DBPC RCT in AR (eg, treatment satisfaction questionnaires and illness perception questionnaires).<sup>43</sup> The selected publications were rated as to whether they included assessments of construct validity (convergent, discriminant, cross-sectional, longitudinal, or language validity), content validity (the comprehensiveness of item selection), criterion validity (concurrent or predictive vs a criterion standard or an outcome), reliability (inter-/intrarater, test-retest, or internal consistency), and responsiveness (ability to detect change, including the MCID), defined according to COSMIN.<sup>32-35</sup>

## RESULTS AND DISCUSSION

### Search results

Our search initially yielded a total of 1705 potentially relevant publications (Figure 2). Many of these publications were spurious hits unrelated to AR; they were produced by our decision to use the key word root “rhin\*” (to include “rhinasthma,” for example) rather than “rhini\*.” After screening the initial list of publications and applying our exclusion criteria, we examined 102 full-text publications in detail. We excluded (1) PROs tested exclusively in rhinosinusitis or asthma, (2) tools used to diagnose

AR or estimate the prevalence of AR in population studies (such as the Score for Allergic Rhinitis<sup>44,45</sup>) or classify types of AR (such as the Allergic Rhinitis and its Impact on Asthma classification),<sup>46,47</sup> (3) PROs used to measure satisfaction with treatment or the treatment outcome alone (such as the ESPIA questionnaire),<sup>48</sup> and (4) illness perception questionnaires.<sup>43,49</sup>

After further exclusions, 55 full-text publications were included in the present review (see Table E1 in this article’s Online Repository at [www.jaci-inpractice.org](http://www.jaci-inpractice.org)).

### Symptom scores

Despite the publication of EMA’s observations in 2008, we found that symptom scores have not undergone extensive validation and—with the possible exception of nasal congestion scores—have been far less extensively validated than disease control scores and QOL scores (see Table E1; see also Table I). Hence, EMA’s call for more validation is as relevant now as it was in 2008.<sup>30</sup> This lack of validation contrasts greatly with the high frequency with which total symptom scores (such as the Rhinoconjunctivitis Total Symptom Score [RTSS]) have been used in major DBPC RCTs, and the regulatory authorities’ recommendations of these scores. Indeed, we did not find any specific references to the psychometric validation of the frequently reported RTSS, other than a study of the MCID. Devillier et al<sup>62</sup> used anchor- and SD-based methods to determine the MCID in the RTSS (6 symptoms scored from 0 to 3, giving a total score ranging from 0 to 18) as part of a multicenter study of 806 patients (253 children, 250 adolescents, and 303 adults) consulting for grass pollen–induced AR (see Table E1). The researchers suggested that the MCID could be rounded down to 1 but emphasized that any subsequently determined MCIDs for AR induced by other allergens may not be the same as those determined for grass pollen–induced AR.<sup>62</sup> Despite the

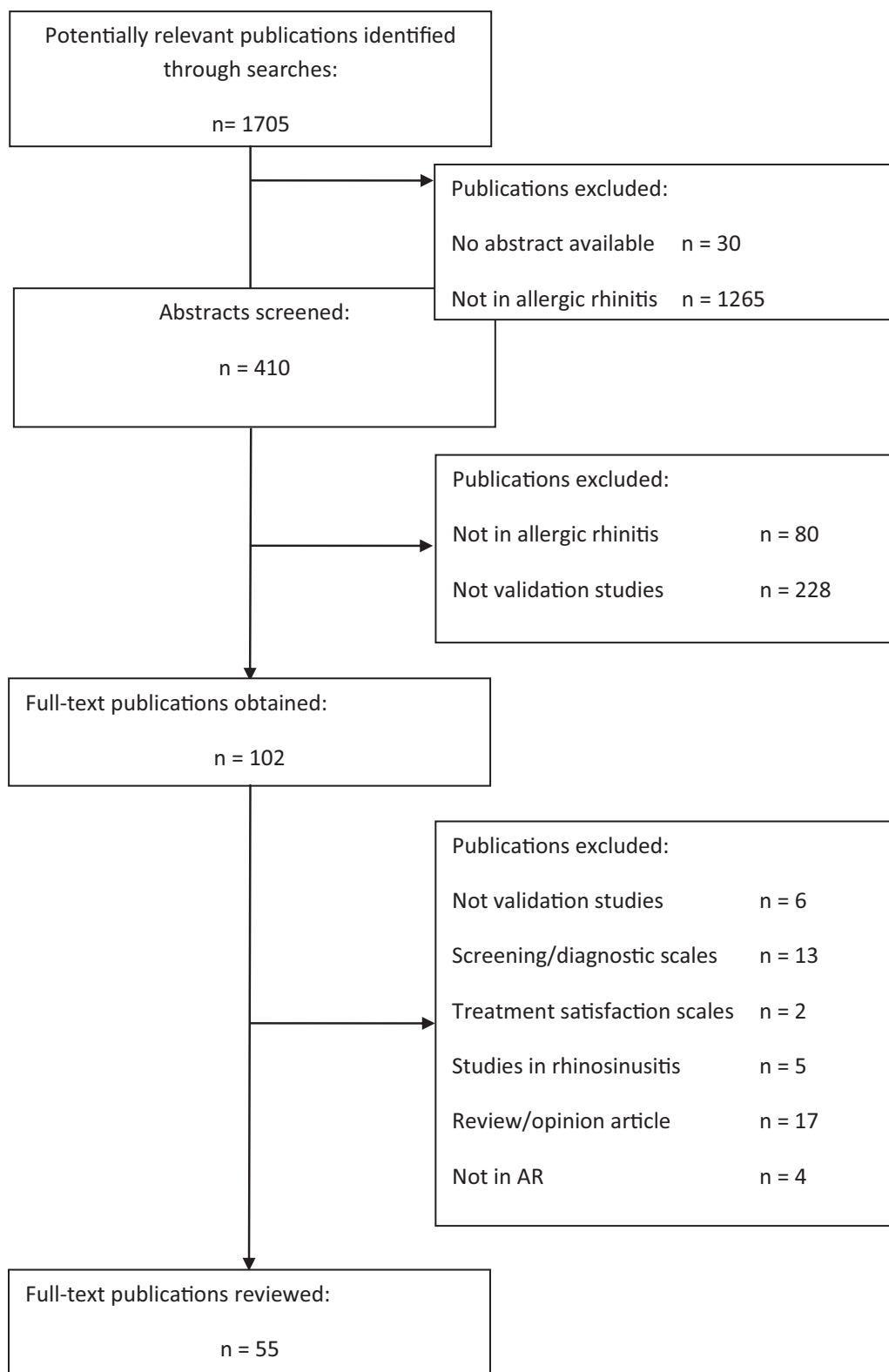


**FIGURE 2.** Study flowchart.

lack of specific psychometric validation studies for the RTSS, other reports have highlighted the strong or moderate correlations between this and other commonly used PROs, such as the VAS. In a further analysis of the aforementioned study

population of 806 patients, Devillier et al<sup>100</sup> showed that weekly changes in RQLQ were correlated with changes in the VAS and RTSS scores. In all 3 age groups, the VAS score was well correlated with the weekly mean RTSS score (Pearson  $r$ , 0.79-

**TABLE I.** Summary of validation status for symptom scores, disease control scores, QOL scores, and other PROs in AR\*

| Name of instrument                                                  | Type of instrument                    | Study population(s)            | References                                                                                                                | Construct validity | Content validity | Criterion validity | Reliability | Responsiveness | MCID (anchor) | MCID (SD) |
|---------------------------------------------------------------------|---------------------------------------|--------------------------------|---------------------------------------------------------------------------------------------------------------------------|--------------------|------------------|--------------------|-------------|----------------|---------------|-----------|
| NOSE                                                                | Symptom score                         | Children and adults            | Stewart et al, <sup>50</sup><br>Zicari et al <sup>51</sup>                                                                | ✓                  | ✓                | ✓                  | ✓           | ✓              | —             | —         |
| CQ5 and CQ7                                                         | Symptom score                         | Adults                         | Stull et al, <sup>52,53</sup><br>Valero et al <sup>54</sup>                                                               | ✓                  | ✓                | —                  | ✓           | ✓              | —             | —         |
| RASS                                                                | Symptom score                         | Adults                         | Wasserfallen et al <sup>55</sup>                                                                                          | ✓                  | ✓                | —                  | ✓           | ✓              | —             | —         |
| NtSS                                                                | Symptom score                         | Teens and adults               | Santanello et al <sup>56</sup>                                                                                            | ✓                  | —                | —                  | ✓           | ✓              | —             | —         |
| RSUI                                                                | Symptom score                         | Adults                         | Revicki et al, <sup>57</sup> Lo et al <sup>58</sup>                                                                       | ✓                  | —                | ✓                  | ✓           | —              | —             | —         |
| PBI-AR                                                              | Symptom score                         | Adults                         | Franzke et al, <sup>59</sup><br>Demoly et al <sup>60</sup>                                                                | ✓                  | —                | ✓                  | ✓           | —              | —             | ✓         |
| SoAR                                                                | Symptom score                         | Adults                         | Johnson et al <sup>61</sup>                                                                                               | ✓                  | —                | —                  | ✓           | —              | —             | —         |
| RTSS                                                                | Symptom score                         | Children, teens,<br>and adults | Devillier et al <sup>62</sup>                                                                                             | —                  | —                | —                  | —           | —              | ✓             | ✓         |
| AdSS                                                                | Symptom score                         | Children, teens,<br>and adults | Grouin et al <sup>63</sup>                                                                                                | ✓                  | —                | —                  | —           | —              | —             | —         |
| A T4NSS                                                             | Symptom score                         | Adults                         | Barnes et al <sup>36</sup>                                                                                                | —                  | —                | —                  | —           | —              | ✓             | ✓         |
| A T5SS (in Japanese cedar/cypress pollinosis)                       | Symptom score                         | Adults                         | Higaki et al <sup>64</sup>                                                                                                | —                  | —                | —                  | —           | —              | ✓             | —         |
| A T6SS (in Japanese cedar/cypress pollinosis)                       | Symptom score                         | Adults                         | Higaki et al <sup>64</sup>                                                                                                | —                  | —                | —                  | —           | —              | ✓             | —         |
| ACS                                                                 | Combined symptom and medication score | Teens and adults               | Hafner et al <sup>65,66</sup>                                                                                             | ✓                  | ✓                | —                  | ✓           | —              | —             | —         |
| CARAT10                                                             | Disease control score                 | Adults                         | Fonseca et al, <sup>67,68</sup><br>van der Leeuw et al <sup>69</sup>                                                      | ✓                  | —                | ✓                  | ✓           | ✓              | ✓             | ✓         |
| CARATKids                                                           | Disease control score                 | Children and teens             | Linhares et al, <sup>47</sup><br>Borrego et al, <sup>70</sup><br>Amaral et al, <sup>71</sup><br>Emons et al <sup>72</sup> | ✓                  | ✓                | —                  | ✓           | ✓              | ✓             | ✓         |
| RCAT                                                                | Disease control score                 | Teens and adults               | Schatz et al, <sup>73,74</sup><br>Meltzer et al, <sup>75</sup><br>Fernandes et al <sup>76</sup>                           | ✓                  | ✓                | —                  | ✓           | —              | ✓             | ✓         |
| RCSS                                                                | Disease control score                 | Adults                         | Boulay and Boulet <sup>77</sup>                                                                                           | ✓                  | ✓                | ✓                  | ✓           | —              | —             | —         |
| ARCT                                                                | Disease control score                 | Children, teens,<br>and adults | Demoly et al, <sup>78</sup><br>Wang et al, <sup>79,80</sup><br>Zhu et al <sup>81</sup>                                    | ✓                  | ✓                | —                  | ✓           | ✓              | —             | —         |
| VAS for the burden of AR in primary care                            | Disease control score                 | Adults                         | Demoly et al <sup>82</sup>                                                                                                | ✓                  | —                | —                  | —           | ✓              | —             | —         |
| VASs for the control of overall, nasal, ocular, and asthma symptoms | Disease control score                 | Not stated                     | Caimmi et al <sup>83</sup>                                                                                                | —                  | —                | —                  | ✓           | —              | —             | —         |

|            |           |                    |                                                                                           |   |   |   |   |   |   |   |
|------------|-----------|--------------------|-------------------------------------------------------------------------------------------|---|---|---|---|---|---|---|
| RQLQ       | QOL score | Adults             | Leong et al, <sup>84,85</sup><br>Okuda et al, <sup>86</sup><br>Yüksel et al <sup>87</sup> | ✓ | ✓ | — | ✓ | ✓ | — | — |
| Mini-RQLQ  | QOL score | Adults             | Barnes et al, <sup>36</sup><br>Schatz et al, <sup>74</sup><br>Yüksel et al <sup>87</sup>  | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| RAPP       | QOL score | Teens and adults   | Molinengo et al, <sup>88</sup><br>Braidó et al <sup>89</sup>                              | ✓ | ✓ | ✓ | ✓ | ✓ | — | ✓ |
| AAQQ       | QOL score | Adults             | AbuRuz et al <sup>90</sup>                                                                | ✓ | ✓ | — | ✓ | ✓ | ✓ | — |
| PADQLQ     | QOL score | Children and teens | Roberts et al, <sup>91</sup><br>Kiotseridis et al <sup>92</sup>                           | ✓ | — | — | ✓ | ✓ | ✓ | ✓ |
| ESPRINT    | QOL score | Adults             | ESPRINT Study<br>Investigators <sup>93</sup>                                              | ✓ | ✓ | — | ✓ | ✓ | — | — |
| RQLQ(S)    | QOL score | Adults             | Juniper et al <sup>94</sup>                                                               | ✓ | — | — | ✓ | ✓ | ✓ | — |
| RHINASTHMA | QOL score | Adults             | Baiardini et al <sup>95</sup>                                                             | ✓ | — | ✓ | ✓ | ✓ | — | — |
| JRQLQ      | QOL score | Adults             | Higaki et al, <sup>64</sup><br>Okuda et al <sup>86</sup>                                  | ✓ | — | — | ✓ | — | ✓ | — |
| NRQLQ      | QOL score | Adults             | Juniper et al <sup>96</sup>                                                               | ✓ | — | — | ✓ | ✓ | — | — |
| PRQLQ      | QOL score | Children and teens | Juniper et al, <sup>97</sup><br>Yüksel et al <sup>98</sup>                                | ✓ | — | — | ✓ | ✓ | — | — |
| Ped-AR-QOL | QOL score | Children and teens | Mavroudi et al <sup>99</sup>                                                              | ✓ | — | — | ✓ | — | — | — |

AAQQ, Arabic Allergic Rhinitis Quality of Life Questionnaire; AdSS, adjusted symptom score; ESPRINT, Cuestionario Español de Calidad de Vida en RINiTis; JRQLQ, Japan Rhinoconjunctivitis Quality of Life Questionnaire; MID, minimal important difference; NRQLQ, Nocturnal Rhinoconjunctivitis Quality of Life Questionnaire; NtSS, nighttime symptoms score; PADQLQ, Pediatric Allergic Disease Quality of Life Questionnaire; PBI-AR, Patient Benefit Index—Allergic Rhinitis; Ped-AR-QOL, Pediatric Allergic Rhinitis Quality of Life Questionnaire; PRQLQ, Paediatric Rhinoconjunctivitis Quality of Life Questionnaire; RAPP, RhinAsthma Patient Perspective; RASS, rhinoconjunctivitis and asthma symptom score; RHINASTHMA, a disease-specific QOL questionnaire; RQLQ, Rhinoconjunctivitis Quality of Life Questionnaire; RQLQ(S), Standardized Rhinoconjunctivitis Quality of Life Questionnaire; RSUI, Rhinitis Symptom Utility Index; SoAR, Scale of Allergic Rhinitis; T4NSS, a 4-item total nasal symptom score.

Note. Age classes of study subjects are defined as follows: children, younger than 13 y; teens, aged 13-17 y; adults, aged 18 y and older.

\*This is on the basis of the publications listed in Table E1 in this article's Online Repository at [www.jaci-inpractice.org](http://www.jaci-inpractice.org).

**TABLE II.** Strengths and weaknesses of the various types of PROs assessed for validation status

| Type of score                          | Strengths                                                                                                                                                                                                                                                                                                           | Weaknesses                                                                                                                                                                                                                                                       |
|----------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Total symptom scores                   | <ul style="list-style-type: none"> <li>• Familiar to clinicians and researchers, because of its extensive use in the literature</li> <li>• The PRO that corresponds closely to a “physiological” score, reflecting the effect of disease on the body</li> <li>• Can be scored rapidly once or more a day</li> </ul> | <ul style="list-style-type: none"> <li>• Not extensively validated</li> <li>• Subject to interindividual differences in the perception of symptom severity</li> </ul>                                                                                            |
| Nasal congestion scores                | <ul style="list-style-type: none"> <li>• Extensively validated</li> <li>• The PRO that corresponds closely to a “physiological” score, reflecting the effect of disease on the body</li> <li>• Can be scored rapidly once or more a day</li> </ul>                                                                  | <ul style="list-style-type: none"> <li>• Does not reflect all the symptoms of AR</li> <li>• Subject to interindividual differences in the perception of symptom severity</li> </ul>                                                                              |
| Combined symptom and medication scores | <ul style="list-style-type: none"> <li>• Conceptually attractive, because the use of rescue medical testifies to a need for symptom relief</li> <li>• Can be scored rapidly once or more a day</li> </ul>                                                                                                           | <ul style="list-style-type: none"> <li>• Not extensively validated</li> <li>• Arbitrary balance between the “symptom” and the “medication” parts of the score</li> <li>• Subject to interindividual differences in the perception of symptom severity</li> </ul> |
| Disease control scores                 | <ul style="list-style-type: none"> <li>• Extensively validated</li> <li>• Takes account of interindividual differences in the perception of symptom severity</li> <li>• May correspond most closely to “real-life” efficacy</li> </ul>                                                                              | <ul style="list-style-type: none"> <li>• Relatively long administration time</li> <li>• Poor time resolution (retrospective/reflective scoring over weeks, typically)</li> </ul>                                                                                 |
| QOL scores                             | <ul style="list-style-type: none"> <li>• Extensively validated</li> </ul>                                                                                                                                                                                                                                           | <ul style="list-style-type: none"> <li>• Relatively long administration time</li> <li>• Far removed from a clinical effect</li> <li>• Poor time resolution (retrospective/reflective scoring over weeks, typically)</li> </ul>                                   |

0.88) and moderately correlated with the weekly mean RQLQ score (Pearson  $r$ , 0.64–0.80); in other words, the study by Devillier et al can be interpreted as an assessment of construct validity for the RTSS.

Higaki et al<sup>64</sup> calculated the MCID for a 6-item symptom score (T6SS) that had been extracted from data on the Japan RQLQ and comprised the same 6 individual symptom scores as the RTSS (sneezing, rhinorrhoea, nasal congestion, itchy nose, itchy eyes, and watery eyes).<sup>64</sup> The data came from a study of 55 patients with Japanese cedar/cypress pollinosis (17 males, 38 females; mean age, 53.1 years<sup>23–29</sup>). The Japan RQLQ rates each individual symptom score from 0 (no symptoms) to 4 (very severe symptoms), whereas the RTSS rates each score from 0 to 3. Furthermore, clinical change was rated against a 5-point “face scale” as the anchor, rather than the commonly used 15-point global rating of change scale. In the 2009 and 2010 Japanese cedar and cypress pollen seasons, the MCID was 0.686 and 0.531 units per item, respectively. Even after considering the differences in the rating scales, these values are markedly higher than those calculated by Devillier et al<sup>62</sup> for the RTSS (ie, ~0.167 per item) in a European population of patients with grass pollen–induced AR.<sup>62</sup> Higaki et al<sup>64</sup> also calculated the MCID for a 5-item symptom score (T5SS) with 3 nasal symptoms (sneezing, rhinorrhoea, and nasal congestion) and 2 ocular symptoms (watery and itchy eyes) rated on a 5-point scale. The per-item MCID was much lower for the T5SS than for the T6SS (0.285 in 2009 and 0.288 in 2010).

We found that the most extensively validated symptom-related scores are those for the Nasal Obstruction Symptom Evaluation (NOSE) scale and the Congestion Quantifier 5-Item Screener (CQ5) and the Congestion Quantifier 7-Item Test

(CQ7).<sup>50,52,53</sup> The NOSE scale is a short, easy-to-complete PRO. Although the NOSE scale’s name refers to “symptom evaluation,” it might be considered as a measure of disease control because the patient is asked “Over the past 1 month, how much of a problem were the following conditions for you?” The NOSE scale also includes a VAS of nasal obstruction. It was assessed for construct, content, and criterion validity, reliability, and responsiveness in patients presenting with drug-refractory symptoms of chronic nasal obstruction for at least 3 months, rather than in a context of AR<sup>50</sup> (Table I). Despite its extensive validation, the NOSE scale score is essentially an individual symptom score with regard to AR, and is probably not a broad-enough measure to be of value in clinical trials in this field.

According to its developers, the CQ was originally formulated as a screening tool for nasal congestion but can distinguish between patients with different levels of severity of AR symptoms, and between cases and controls.<sup>52</sup> It is also sensitive to changes in a patient’s condition. The CQ7 was assessed for construct and content validity, reliability, and responsiveness in a study of 129 patients with confirmed AR, 155 patients with nasal congestion, and 70 control subjects recruited at a total of 49 sites<sup>52</sup> (Table I). Although the CQ focuses on the preceding week (rather than the preceding or current day), we consider that this tool might be a suitable measure of clinical efficacy over a long-term RCT in AR.

### Medication scores

We did not find any publications describing the validation of a “pure” medication score. Again, this complicates the regulatory authorities’ preference for a CSMS as the recommended score in pivotal clinical trials. Despite the presence of the word “control” in its name, the Allergy-Control-SCORE (ACS) is a CSMS.<sup>65,66</sup>

The medication score component is based on a list of 745 commercially available preparations of drugs with relevance in AR and allergic asthma, each of which is attributed a target organ (nose, eyes, or lungs) and a number of points. In a study of 81 patients with AR and/or allergic asthma and 40 healthy controls, the ACS was assessed for construct and content validity, and reliability with regard to nasal, ocular, and lung symptoms<sup>65</sup> and then with regard to nasal and ocular symptoms only<sup>66</sup> (Table I).

### Disease control scores

We found that disease control scales have been extensively validated in AR (see Table E1). However, one should bear in mind that there is no single definition of “disease control”; nevertheless, PROs referred to as disease control scales usually assess most or all of the following domains, albeit to varying extents: (1) symptom intensity and/or frequency (ie, symptom scoring), (2) physical impairments (such as air flow), (3) the frequency and consequences of exacerbations, and (4) the disease burden (ie, the “bothersomeness” of symptoms, and disease-related impairments in social, physical, or professional activities).<sup>42</sup> The Rhinitis Control Scoring System (RCSS) rates 5 symptoms (sneezing, rhinorrhoea, nasal obstruction, nasal itching, and conjunctivitis) in terms of severity and frequency over the previous 7 days, but does not probe bothersomeness or exacerbations; despite the presence of “control” in the score’s name, it should probably be classified as a retrospective symptom score.<sup>77</sup> The RCSS has nevertheless been thoroughly validated (construct, content, and possibly criterion validity, reliability, and internal consistency) in a study of 50 patients with AR, and warrants further investigation as an efficacy criterion.

We found that the most extensively validated disease control PROs were the Rhinitis Control Assessment Test (RCAT) and the Control of Allergic Rhinitis and Asthma Test (CARAT) (see Table E1; see also Table I). In a study of data from the placebo arm of 2 DBPC RCTs (on 29 grass pollen–allergic patients and 76 birch pollen–allergic patients), Liedtke et al<sup>101</sup> found strong correlations between the RCAT and RQLQ scores in grass pollen–allergic and birch pollen–allergic patients ( $r = -0.871$  and  $-0.795$ , respectively), a strong correlation between the RCAT and RTSS scores in grass pollen–allergic patients ( $r = -0.811$ ), and a moderate correlation between the RCAT and RTSS scores in birch pollen–allergic patients ( $r = -0.539$ ).<sup>101</sup> These findings suggest that efficacy in clinical trials may be evidenced by disease control scores.

The CARAT was initially developed as a 17-question test in Portuguese, by using a comprehensive set of methodological steps (including expert and patient panels) to verify the design quality, face validity, and content validity.<sup>102</sup> A 10-item version was then produced (using descriptive analysis, exploratory factor analysis, and internal consistency) and validated in a cross-sectional study of 193 adults with AR and allergic asthma from 15 outpatient clinics in Portugal.<sup>67</sup> On the basis of the studies reviewed here, the CARAT10 has been evaluated for construct and criterion validity, reliability, responsiveness, and the MCID (Table I). Furthermore, the pediatric version of the CARAT (CARATKids) has been evaluated for construct and content validity, reliability, responsiveness, and the MCID (Table I). Another extensively validated disease control score is the Allergic Rhinitis Control Test (ARCT), for which construct and content validity, reliability, and responsiveness to change have been examined.<sup>78–81</sup> The ARCT has notably been applied to the

management of “step-up” and “step-down” approaches to the prescription of symptomatic medications in patients with AR, as a function of the level of disease control achieved.<sup>79,81</sup>

### QOL scales

Health-related QOL scales (and especially the RQLQ and the mini-RQLQ) have been extensively validated in AR, notably because of the pioneering work by RQLQ’s developers. Indeed, the mini-RQLQ and the RQLQ have been considered by some investigators as criterion standards in AR against which the criterion validity of other PROs can be assessed.<sup>60,103</sup>

The RQLQ has been used as a secondary efficacy criterion in large, multicenter trials of sublingual grass pollen AIT formulations, where changes in the QOL were correlated with changes in a symptom score (RTSS).<sup>104,105</sup> However, the use of the RQLQ as a primary end point to be scored daily in a pivotal DBPC RCT in AR is compromised by the relatively long administration time (about 10 minutes). Furthermore, the recall period in the RQLQ is 7 days, and so this PRO is not an instantaneous score. Furthermore, the use of instruments with recall periods of 1 week or more may mean that some relevant days with symptoms will be missed.<sup>29</sup> The mini-RQLQ is a more rapid alternative and has also been extensively validated. Several other AR-related QOL tools have been partially or fully validated in specific contexts and languages; these include the RHINASTHMA disease-specific QOL questionnaire, the RhinAsthma Patient Perspective (derived from the RHINASTHMA questionnaire), the Arabic Allergic Rhinitis Quality of Life Questionnaire, the Pediatric Allergic Disease Quality of Life Questionnaire, and the Cuestionario Español de Calidad de Vida en RINiTis questionnaire.<sup>88,90,91,93,95</sup>

### Other scores and perspectives for future research

Finally, we decided to review the Allergy Outcome Survey (AOS).<sup>106</sup> This is something of a hybrid score, because it combines items related to disease control, symptom severity, medication use, and treatment outcome. Kemker et al<sup>106</sup> concluded that the AOS was a brief, reliable, easy-to-use performance metric that could evaluate change with therapy when used alone or combined with other PROs.<sup>106</sup> However, AOS’s feasibility as a regularly scored PRO in an RCT is debatable. It might also be of value to explore the wider use of illness perception scales (such as the Brief Illness Perception Questionnaire) and treatment satisfaction questionnaires<sup>43,48</sup>—at least as secondary end points. On the basis of our assessments of validation status, the strengths and weaknesses of the various types of scores are presented in Table II.

As mentioned earlier, the present systematic review was prompted by regulatory calls for a validated symptom score that could be used in pivotal clinical trials. By definition, PROs are reported by patients, and patients spend most of their time away from medical facilities. Hence, it would be particularly interesting to explore the combination of a validated PRO in AR with the “m-health” and “e-health” information technology tools that are now widely available among the general public (eg, smartphones with Internet access). Many different apps have been developed for patients with asthma, although the focus is usually self-management and not PRO data collection *per se*.<sup>107,108</sup> Some research in the field of AR has already been described. A recent review identified 72 apps related to AR; although most were related to patient education and pollen count, 14 (19.4%) were

related to symptom tracking.<sup>109</sup> By way of an example, the Mobile Airways Sentinel Network rhinitis tool is used to implement care pathways for AR by using a smartphone- or tablet-based application and a clinical decision support system.<sup>83,110,111</sup> The user self-evaluates his or her level of AR control on a VAS. The system's developers consider that the Mobile Airways Sentinel Network rhinitis VAS app (currently available in 16 languages and 23 countries) is a reliable, valid tool for assessing AR allergic control at the population level.<sup>110</sup>

## CONCLUSIONS

Despite the regulatory authorities' current emphasis on CSMSs as the basis for evaluating the efficacy of AIT in AR, neither of the latter's 2 components (the symptom score or the medication score) has been extensively validated. The most frequently used symptom score (RTSS) has not been thoroughly validated, although the MCID has been determined in patients with grass pollen-induced AR. The validity of medication scores in AR has not been studied. In contrast, disease control scales and health-related QOL scales have been extensively validated. However, the adoption of a disease control or QOL scale as a primary end point in clinical trials will require a paradigm shift within the clinical community and regulatory authorities. The regulatory authorities' current requirement for daily symptom score data makes it difficult to envisage the use of retrospective or reflective disease control scores or disease-related QOL scores as primary efficacy end points. Nevertheless, some experts consider that a paradigm shift in this field is required, that is, with further emphasis on the long-term control (over several months or years) of atopic diseases rather than on symptom control during a clinical trial that sometimes may last for just a few weeks or months (eg, a single pollen season).<sup>112</sup> In the field of asthma, the Asthma Control Questionnaire<sup>113</sup> and the Asthma Control Test<sup>114</sup> have recently been recommended for use in both primary care and other clinical settings.<sup>115</sup>

We have the following closing comments and recommendations: (1) the RTSS and a CSMS should undergo full psychometric validation (notably with regard to construct, content, and criterion validity) in a well-powered clinical trial in AR as a matter of some urgency; (2) in collaboration with all the relevant stakeholders, the regulatory authorities should evaluate the possible use of a disease control score as the primary efficacy criterion in a phase III clinical trial; (3) whenever possible, a phase III clinical trial in AR should have a QOL score as a secondary efficacy criterion; and (4) PROs validated in AR should be combined with m-health and e-health tools for symptom/disease control tracking.

## REFERENCES

1. Blaiss MS, Hammerby E, Robinson S, Kennedy-Martin T, Buchs S. The burden of allergic rhinitis and allergic rhinoconjunctivitis in adolescents: a literature review. *Ann Allergy Asthma Immunol* 2018;121:43-52.e3.
2. Linneberg A, Dam Petersen K, Hahn-Pedersen J, Hammerby E, Serup-Hansen N, Boxall N. Burden of allergic respiratory disease: a systematic review. *Clin Mol Allergy* 2016;14:12.
3. Ozdoganoglu T, Songu M. The burden of allergic rhinitis and asthma. *Ther Adv Respir Dis* 2012;6:11-23.
4. Maspero J, Lee BW, Katelaris CH, Potter PC, Cingi C, Lopatin A, et al. Quality of life and control of allergic rhinitis in patients from regions beyond western Europe and the United States. *Clin Exp Allergy* 2012;42:1684-96.
5. Demoly P, Maigret P, Elias Billon I, Allaert FA. Allergic rhinitis increases the risk of driving accidents. *J Allergy Clin Immunol* 2017;140:614-6.
6. Shaker M. New insights into the allergic march. *Curr Opin Pediatr* 2014;26:516-20.
7. Thomsen SF. Epidemiology and natural history of atopic diseases. *Eur Clin Respir J* 2015;2:24642.
8. Ridolo E, Montagni M, Melli V, Braidò F, Incorvaia C, Canonica GW. Pharmacotherapy of allergic rhinitis: current options and future perspectives. *Expert Opin Pharmacother* 2014;15:73-83.
9. Porcaro F, Corsello G, Pajno GB. SLIT's prevention of the allergic march. *Curr Allergy Asthma Rep* 2018;18:31.
10. Pfaar O, Alvaro M, Cardona V, Hamelmann E, Mosges R, Kleine-Tebbe J. Clinical trials in allergen immunotherapy: current concepts and future needs. *Allergy* 2018;73:1775-83.
11. Dhami S, Nurmatov U, Arasi S, Khan T, Asaria M, Zaman H, et al. Allergen immunotherapy for allergic rhinoconjunctivitis: a systematic review and meta-analysis. *Allergy* 2017;72:1597-631.
12. Jutel M, Agache I, Bonini S, Burks AW, Calderon M, Canonica W, et al. International consensus on allergen immunotherapy, II: mechanisms, standardization, and pharmacoeconomics. *J Allergy Clin Immunol* 2016;137:358-68.
13. Roberts G, Pfaar O, Akdis CA, Ansotegui IJ, Durham SR, Gerth van Wijk R, et al. EAACI guidelines on allergen immunotherapy: allergic rhinoconjunctivitis. *Allergy* 2018;73:765-98.
14. Jutel M, Agache I, Bonini S, Burks AW, Calderon M, Canonica W, et al. International consensus on allergy immunotherapy. *J Allergy Clin Immunol* 2015;136:556-68.
15. Martignago I, Incorvaia C, Ridolo E. Preventive actions of allergen immunotherapy: the facts and the effects in search of evidence. *Clin Mol Allergy* 2017;15:13.
16. Pfaar O, Bonini S, Cardona V, Demoly P, Jakob T, Jutel M, et al. Perspectives in allergen immunotherapy: 2017 and beyond. *Allergy* 2018;73:5-23.
17. Burks AW, Calderon MA, Casale T, Cox L, Demoly P, Jutel M, et al. Update on allergy immunotherapy: American Academy of Allergy, Asthma & Immunology/European Academy of Allergy and Clinical Immunology/PRACTALL consensus report. *J Allergy Clin Immunol* 2013;131:1288-1296.e3.
18. Bonertz A, Roberts GC, Hoefnagel M, Timon M, Slater JE, Rabin RL, et al. Challenges in the implementation of EAACI guidelines on allergen immunotherapy: a global perspective on the regulation of allergen products. *Allergy* 2018;73:64-76.
19. Baiardini I, Bousquet PJ, Brzoza Z, Canonica GW, Compalati E, Flocchi A, et al. Recommendations for assessing patient-reported outcomes and health-related quality of life in clinical trials on allergy: a GA(2)LEN taskforce position paper. *Allergy* 2010;65:290-5.
20. Baiardini I, Braidò F, Bindeslev-Jensen C, Bousquet PJ, Brzoza Z, Canonica GW, et al. Recommendations for assessing patient-reported outcomes and health-related quality of life in patients with urticaria: a GA(2) LEN taskforce position paper. *Allergy* 2011;66:840-4.
21. Brozek JL, Bousquet J, Agache I, Agarwal A, Bachert C, Bosnic-Anticevich S, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines—2016 revision. *J Allergy Clin Immunol* 2017;140:950-8.
22. Calderon MA, Casale TB, Togias A, Bousquet J, Durham SR, Demoly P. Allergen-specific immunotherapy for respiratory allergies: from meta-analysis to registration and beyond. *J Allergy Clin Immunol* 2011;127:30-8.
23. Calderon MA, Kleine-Tebbe J, Linneberg A, De Blay F, Hernandez Fernandez de Rojas D, Virchow JC, et al. House dust mite respiratory allergy: an overview of current therapeutic strategies. *J Allergy Clin Immunol Pract* 2015;3:843-55.
24. Canonica GW, Baena-Cagnani CE, Bousquet J, Bousquet PJ, Lockey RF, Mallon HJ, et al. Recommendations for standardization of clinical trials with allergen specific immunotherapy for respiratory allergy: a statement of a World Allergy Organization (WAO) taskforce. *Allergy* 2007;62:317-24.
25. Nelson HS, Calderon MA, Bernstein DI, Casale TB, Durham SR, Andersen JS, et al. Allergen immunotherapy clinical trial outcomes and design: working toward harmonization of methods and principles. *Curr Allergy Asthma Rep* 2017;17:18.
26. Papadopoulos NG, Agache I, Bavbek S, Bilo BM, Braidò F, Cardona V, et al. Research needs in allergy: an EAACI position paper, in collaboration with EFA. *Clin Transl Allergy* 2012;2:21.
27. Passalacqua G. Recommendations for appropriate sublingual immunotherapy clinical trials. *World Allergy Organ J* 2014;7:21.
28. Pfaar O, Anders C, Klimek L. Clinical outcome measures of specific immunotherapy. *Curr Opin Allergy Clin Immunol* 2009;9:208-13.

29. Pfaar O, Demoly P, Gerth van Wijk R, Bonini S, Bousquet J, Canonica GW, et al. Recommendations for the standardization of clinical outcomes used in allergen immunotherapy trials for allergic rhinoconjunctivitis: an EAACI position paper. *Allergy* 2014;69:854-67.
30. European Medicines Agency. European Medicines Agency (EMA) guideline on the clinical development of products for specific immunotherapy for the treatment of allergic diseases. 2008. Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500003605.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003605.pdf). Accessed January 12, 2019.
31. US Department of Health and Human Services; Food and Drug Administration; Center for Drug Evaluation and Research. Allergic rhinitis: developing drug products for treatment: guidance for industry (draft guidance). 2016. Available from: <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071293.pdf>. Accessed January 12, 2019.
32. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement Instruments. *BMC Med Res Methodol* 2006;6:2.
33. Mokkink LB, Prinsen CA, Bouter LM, Vet HC, Terwee CB. The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016;20:105-13.
34. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171-9.
35. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147-57.
36. Barnes ML, Vaidyanathan S, Williamson PA, Lipworth BJ. The minimal clinically important difference in allergic rhinitis. *Clin Exp Allergy* 2010;40:242-50.
37. Meltzer EO, Wallace D, Dykewicz M, Shneyer L. Minimal clinically important difference (MCID) in allergic rhinitis: Agency for Healthcare Research and Quality or anchor-based thresholds? *J Allergy Clin Immunol Pract* 2016;4:682-688.e6.
38. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102-9.
39. Juniper EF, Guyatt GH. Development and testing of a new measure of health status for clinical trials in rhinoconjunctivitis. *Clin Exp Allergy* 1991;21:77-83.
40. Juniper EF, Guyatt GH, Griffith LE, Ferrie PJ. Interpretation of rhinoconjunctivitis quality of life questionnaire data. *J Allergy Clin Immunol* 1996;98:843-5.
41. Dietz de Loos DA, Segboer CL, Gevorgyan A, Fokkens WJ. Disease-specific quality-of-life questionnaires in rhinitis and rhinosinusitis: review and evaluation. *Curr Allergy Asthma Rep* 2013;13:162-70.
42. Demoly P, Calderon MA, Casale T, Scadding G, Annesi-Maesano I, Braun JJ, et al. Assessment of disease control in allergic rhinitis. *Clin Transl Allergy* 2013;3:7.
43. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res* 2006;60:631-7.
44. Annesi-Maesano I, Didier A, Klossek M, Chantal I, Moreau D, Bousquet J. The score for allergic rhinitis (SFAR): a simple and valid assessment method in population studies. *Allergy* 2002;57:107-14.
45. Lam SC, Yeung CCY, Chan JHM, Lam DWC, Lam AHY, Annesi-Maesano I, et al. Adaptation of the score for allergic rhinitis in the Chinese population: psychometric properties and diagnostic accuracy. *Int Arch Allergy Immunol* 2017;173:213-24.
46. Jauregui I, Davila I, Sastre J, Bartra J, del Cuvillo A, Ferrer M, et al. Validation of ARIA (Allergic Rhinitis and its Impact on Asthma) classification in a pediatric population: the PEDRIAL study. *Pediatr Allergy Immunol* 2011;22:388-92.
47. Linhares DV, da Fonseca JA, Borrego LM, Matos A, Pereira AM, Sa-Sousa A, et al. Validation of control of allergic rhinitis and asthma test for children (CARATKids)—a prospective multicenter study. *Pediatr Allergy Immunol* 2014;25:173-9.
48. Justicia JL, Cardona V, Guardia P, Ojeda P, Olaguibel JM, Vega JM, et al. Validation of the first treatment-specific questionnaire for the assessment of patient satisfaction with allergen-specific immunotherapy in allergic patients: the ESPIA questionnaire. *J Allergy Clin Immunol* 2013;131:1539-46.
49. Pesut D, Raskovic S, Tomic-Spiric V, Bulajic M, Bogic M, Bursuc B, et al. Gender differences revealed by the Brief Illness Perception Questionnaire in allergic rhinitis. *Clin Respir J* 2014;8:364-8.
50. Stewart MG, Witsell DL, Smith TL, Weaver EM, Yueh B, Hannley MT. Development and validation of the Nasal Obstruction Symptom Evaluation (NOSE) scale. *Otolaryngol Head Neck Surg* 2004;130:157-63.
51. Zicari AM, Occasi F, Montanari G, Indinnimeo L, De Castro G, Tancredi G, et al. Intranasal budesonide in children affected by persistent allergic rhinitis and its effect on nasal patency and Nasal Obstruction Symptom Evaluation (NOSE) score. *Curr Med Res Opin* 2015;31:391-6.
52. Stull DE, Krouse J, Meltzer EO, Roberts L, Kim S, Frank L, et al. Development and validation of the Congestion Quantifier seven-item test (CQ7): a screening tool for nasal congestion. *Value Health* 2007;10:457-65.
53. Stull DE, Meltzer EO, Krouse JH, Roberts L, Kim S, Frank L, et al. The congestion quantifier five-item test for nasal congestion: refinement of the congestion quantifier seven-item test. *Am J Rhinol Allergy* 2010;24:34-8.
54. Valero A, Mullol J, Herdman M, Rosales MJ, Spanish CQ7 Study Group. Measuring outcomes in allergic rhinitis: psychometric characteristics of a Spanish version of the congestion quantifier seven-item test (CQ7). *Health Qual Life Outcomes* 2011;9:14.
55. Wasserfallen JB, Gold K, Schulman KA, Baraniuk JN. Development and validation of a rhinoconjunctivitis and asthma symptom score for use as an outcome measure in clinical trials. *J Allergy Clin Immunol* 1997;100:16-22.
56. Santanello NC, DeMuro-Mercon C, Shah SR, Schenkel EJ, Ratner PH, Dass SB, et al. Validation of the nighttime symptoms score as a clinically relevant measure of allergic rhinitis. *Allergy Asthma Proc* 2006;27:231-9.
57. Revicki DA, Leidy NK, Brennan-Diemer F, Thompson C, Toggias A. Development and preliminary validation of the multiattribute Rhinitis Symptom Utility Index. *Qual Life Res* 1998;7:693-702.
58. Lo PS, Tong MC, Revicki DA, Lee CC, Woo JK, Lam HC, et al. Rhinitis Symptom Utility Index (RSUI) in Chinese subjects: a multiattribute patient-preference approach. *Qual Life Res* 2006;15:877-87.
59. Francke N, Schafer I, Jost K, Blome C, Rustenbach SJ, Reich K, et al. A new instrument for the assessment of patient-defined benefit in the treatment of allergic rhinitis. *Allergy* 2011;66:665-70.
60. Demoly P, Aubier M, de Blay F, Wessel F, Clerson P, Maigret P. Evaluation of patients' expectations and benefits in the treatment of allergic rhinitis with a new tool: the patient benefit index—the benefica study. *Allergy Asthma Clin Immunol* 2015;11:8.
61. Johnson B, Harrington R, Perz C. Validation of the scale of allergic rhinitis. *Psychol Health Med* 2004;9:217-25.
62. Devillier P, Chassany O, Vicaut E, de Beaumont O, Robin B, Dreyfus JF, et al. The minimally important difference in the Rhinoconjunctivitis Total Symptom Score in grass-pollen-induced allergic rhinoconjunctivitis. *Allergy* 2014;69:1689-95.
63. Grouin JM, Vicaut E, Jean-Alphonse S, Demoly P, Wahn U, Didier A, et al. The average Adjusted Symptom Score, a new primary efficacy end-point for specific allergen immunotherapy trials. *Clin Exp Allergy* 2011;41:1282-8.
64. Higaki T, Okano M, Kariya S, Fujiwara T, Haruna T, Hirai H, et al. Determining minimal clinically important differences in Japanese cedar/cypress pollinosis patients. *Allergol Int* 2013;62:487-93.
65. Hafner D, Reich K, Matricardi PM, Meyer H, Kettner J, Narkus A. Prospective validation of "Allergy-Control-SCORE™": a novel symptom-medication score for clinical trials. *Allergy* 2011;66:629-36.
66. Hafner D, Reich K, Zschocke I, Lotzin A, Meyer H, Kettner J, et al. Prospective validation of the "rhino conjunctivitis allergy-control-SCORE(c)" (RC-ACS(c)). *Clin Transl Allergy* 2012;2:17.
67. Fonseca JA, Nogueira-Silva L, Morais-Almeida M, Azevedo L, Sa-Sousa A, Branco-Ferreira M, et al. Validation of a questionnaire (CARAT10) to assess rhinitis and asthma in patients with asthma. *Allergy* 2010;65:1042-8.
68. Fonseca JA, Nogueira-Silva L, Morais-Almeida M, Sa-Sousa A, Azevedo LF, Ferreira J, et al. Control of Allergic Rhinitis and Asthma Test (CARAT) can be used to assess individual patients over time. *Clin Transl Allergy* 2012;2:16.
69. van der Leeuw S, van der Molen T, Dekhuijzen PN, Fonseca JA, van Gemert FA, Gerth van Wijk R, et al. The minimal clinically important difference of the Control of Allergic Rhinitis and Asthma Test (CARAT): cross-cultural validation and relation with pollen counts. *NPJ Prim Care Respir Med* 2015;25:14107.
70. Borrego LM, Fonseca JA, Pereira AM, Pinto VR, Linhares D, Morais-Almeida M. Development process and cognitive testing of CARATKids—Control of Allergic Rhinitis and Asthma Test for Children. *BMC Pediatr* 2014;14:34.
71. Amaral R, Carneiro AC, Wandalsen G, Fonseca JA, Sole D. Control of Allergic Rhinitis and Asthma Test for Children (CARATKids): validation in Brazil and cutoff values. *Ann Allergy Asthma Immunol* 2017;118:551-556.e2.
72. Emons JA, Flokstra BM, de Jong C, van der Molen T, Brand HK, Arends NJ, et al. Use of the Control of Allergic Rhinitis and Asthma Test (CARATKids) in children and adolescents: validation in Dutch. *Pediatr Allergy Immunol* 2017;28:185-90.
73. Schatz M, Meltzer EO, Nathan R, Derebery MJ, Mintz M, Stanford RH, et al. Psychometric validation of the rhinitis control assessment test: a brief patient-

- completed instrument for evaluating rhinitis symptom control. *Ann Allergy Asthma Immunol* 2010;104:118-24.
74. Schatz M, Zeiger RS, Chen W, Yang SJ, Stanford RH, Garriss CP. A comparison of the psychometric properties of the Mini-Rhinitis Quality of Life Questionnaire and the Rhinitis Control Assessment Test. *Am J Rhinol Allergy* 2012;26:127-33.
  75. Meltzer EO, Schatz M, Nathan R, Garriss C, Stanford RH, Kosinski M. Reliability, validity, and responsiveness of the Rhinitis Control Assessment Test in patients with rhinitis. *J Allergy Clin Immunol* 2013;131:379-86.
  76. Fernandes PH, Matsumoto F, Sole D, Wandalsen GF. Translation into Portuguese and validation of the Rhinitis Control Assessment Test (RCAT) questionnaire. *Braz J Otorhinolaryngol* 2016;82:674-9.
  77. Boulay ME, Boulet LP. The rhinitis control scoring system: development and validation. *Am J Rhinol Allergy* 2016;30:54-9.
  78. Demoly P, Jankowski R, Chassany O, Bessah Y, Allaert FA. Validation of a self-questionnaire for assessing the control of allergic rhinitis. *Clin Exp Allergy* 2011;41:860-8.
  79. Wang Y, Chen H, Zhu R, Liu G, Huang N, Li W, et al. Allergic Rhinitis Control Test questionnaire-driven stepwise strategy to improve allergic rhinitis control: a prospective study. *Allergy* 2016;71:1612-9.
  80. Wang Y, Zhu R, Liu G, Li W, Chen H, Dures JP, et al. Prevalence of uncontrolled allergic rhinitis in Wuhan, China: a prospective cohort study. *Am J Rhinol Allergy* 2014;28:397-403.
  81. Zhu R, Wang J, Wu Y, Yang Y, Huang N, Yang Y, et al. The Allergic Rhinitis Control Test questionnaire is valuable in guiding step-down pharmacotherapy treatment of allergic rhinitis. *J Allergy Clin Immunol Pract* 2019;7:272-8.
  82. Demoly P, Bousquet PJ, Mesbah K, Bousquet J, Devillier P. Visual analogue scale in patients treated for allergic rhinitis: an observational prospective study in primary care: asthma and rhinitis. *Clin Exp Allergy* 2013;43:881-8.
  83. Caimmi D, Baiz N, Tanno LK, Demoly P, Arnavielhe S, Murray R, et al. Validation of the MASK-rhinitis visual analogue scale on smartphone screens to assess allergic rhinitis control. *Clin Exp Allergy* 2017;47:1526-33.
  84. Leong KP, Chan SP, Tang CY, Yeak SC, Saurajen AS, Mok PK, et al. Quality of life of patients with perennial allergic rhinitis: preliminary validation of the Rhinoconjunctivitis Quality of Life Questionnaire in Singapore. *Asian Pac J Allergy Immunol* 1999;17:163-7.
  85. Leong KP, Yeak SC, Saurajen AS, Mok PK, Earnest A, Siow JK, et al. Why generic and disease-specific quality-of-life instruments should be used together for the evaluation of patients with persistent allergic rhinitis. *Clin Exp Allergy* 2005;35:288-98.
  86. Okuda M, Ohkubo K, Goto M, Okamoto H, Konno A, Baba K, et al. Comparative study of two Japanese rhinoconjunctivitis quality-of-life questionnaires. *Acta Otolaryngol* 2005;125:736-44.
  87. Yuksel H, Yilmaz O, Alkan S, Bayrak Degirmenci P, Kirmaz C. Validity and reliability of Turkish version of rhinitis and mini-rhinitis quality of life questionnaires. *Allergol Immunopathol (Madr)* 2009;37:293-7.
  88. Molinengo G, Baiardini I, Braido F, Loera B. RhinAsthma patient perspective: a Rasch validation study. *J Asthma* 2018;55:119-23.
  89. Braido F, Baiardini I, Stagi E, Scichilone N, Rossi O, Lombardi C, et al. RhinAsthma patient perspective: a short daily asthma and rhinitis QoL assessment. *Allergy* 2012;67:1443-50.
  90. AbuRuz SM, Bulatova NR, Tawalbeh MI. Development and validation of the Arabic allergic rhinitis quality of life questionnaire. *Saudi Med J* 2009;30:1577-83.
  91. Roberts G, Hurley C, Lack G. Development of a quality-of-life assessment for the allergic child or teenager with multisystem allergic disease. *J Allergy Clin Immunol* 2003;111:491-7.
  92. Kiotseridis H, Cilio CM, Bjerrner L, Aurivillius M, Jacobsson H, Tunsater A. Swedish translation and validation of the Pediatric Allergic Disease Quality of Life Questionnaire (PADQLQ). *Acta Paediatr* 2011;100:242-7.
  93. ESPRINT Study Group and Investigators. Valero A, Alonso J, Antepara I, Baro E, Colas C, del Cuvillo A, et al. Development and validation of a new Spanish instrument to measure health-related quality of life in patients with allergic rhinitis: the ESPRINT questionnaire. *Value Health* 2007;10:466-77.
  94. Juniper EF, Thompson AK, Ferrie PJ, Roberts JN. Validation of the standardized version of the Rhinoconjunctivitis Quality of Life Questionnaire. *J Allergy Clin Immunol* 1999;104:364-9.
  95. Baiardini I, Pasquali M, Giardini A, Specchia C, Passalacqua G, Venturi S, et al. Rhinasthma: a new specific QoL questionnaire for patients with rhinitis and asthma. *Allergy* 2003;58:289-94.
  96. Juniper EF, Rohrbach T, Meltzer EO. A questionnaire to measure quality of life in adults with nocturnal allergic rhinoconjunctivitis. *J Allergy Clin Immunol* 2003;111:484-90.
  97. Juniper EF, Howland WC, Roberts NB, Thompson AK, King DR. Measuring quality of life in children with rhinoconjunctivitis. *J Allergy Clin Immunol* 1998;101:163-70.
  98. Yuksel H, Yilmaz O, Sogut A, Eser E. Validation and reliability study of the Turkish version of the Pediatric Rhinitis Quality of Life Questionnaire. *Turk J Pediatr* 2009;51:361-6.
  99. Mavroudi A, Chrysoschoou EA, Boyle RJ, Papastergiopoulos A, Karantaglis N, Karagiannidou A, et al. Validation study of the pediatric allergic rhinitis quality of life questionnaire. *Asian Pac J Allergy Immunol* 2016;34:159-65.
  100. Devillier P, Bousquet PJ, Grassin-Delye S, Salvator H, Demoly P, Bousquet J, et al. Comparison of outcome measures in allergic rhinitis in children, adolescents and adults. *Pediatr Allergy Immunol* 2016;27:375-81.
  101. Liedtke JP, Mandl A, Kother J, Chwierski J, Shah-Hosseini K, Raskopf E, et al. RCAT reflects symptom control and quality of life in allergic rhinoconjunctivitis patients. *Allergy* 2018;73:1101-9.
  102. Nogueira-Silva L, Martins SV, Cruz-Correia R, Azevedo LF, Morais-Almeida M, Bugalho-Almeida A, et al. Control of allergic rhinitis and asthma test—a formal approach to the development of a measuring tool. *Respir Res* 2009;10:52.
  103. Juniper EF, Thompson AK, Ferrie PJ, Roberts JN. Development and validation of the mini Rhinoconjunctivitis Quality of Life Questionnaire. *Clin Exp Allergy* 2000;30:132-40.
  104. Didier A, Mallin HJ, Worm M, Horak F, Jager S, Montagut A, et al. Optimal dose, efficacy, and safety of once-daily sublingual immunotherapy with a 5-grass pollen tablet for seasonal allergic rhinitis. *J Allergy Clin Immunol* 2007;120:1338-45.
  105. Durham SR, Yang WH, Pedersen MR, Johansen N, Rak S. Sublingual immunotherapy with once-daily grass allergen tablets: a randomized controlled trial in seasonal allergic rhinoconjunctivitis. *J Allergy Clin Immunol* 2006;117:802-9.
  106. Kemker BJ, Corey JP, Branca J, Gliklich RE. Development of the allergy outcome survey for allergic rhinitis. *Otolaryngol Head Neck Surg* 1999;121:603-5.
  107. Marcano Belisario JS, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. *Cochrane Database Syst Rev* 2013;CD010013.
  108. Tinschert P, Jakob R, Barata F, Kramer JN, Kowatsch T. The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps. *JMIR Mhealth Uhealth* 2017;5:e113.
  109. Zhou AH, Patel VR, Baredes S, Eloy JA, Hsueh WD. Mobile applications for allergic rhinitis. *Ann Otol Rhinol Laryngol* 2018;127:836-40.
  110. Bousquet J, Arnavielhe S, Bedbrook A, Bewick M, Laune D, Mathieu-Dupas E, et al. MASK 2017: ARIA digitally-enabled, integrated, person-centred care for rhinitis and asthma multimorbidity using real-world-evidence. *Clin Transl Allergy* 2018;8:45.
  111. Bousquet J, Caimmi DP, Bedbrook A, Bewick M, Hellings PW, Devillier P, et al. Pilot study of mobile phone technology in allergic rhinitis in European countries: the MASK-rhinitis study. *Allergy* 2017;72:857-65.
  112. Bieber T, Vieths S, Broich K. New opportunities and challenges in the assessment of drugs for atopic diseases. *Allergy* 2016;71:1662-5.
  113. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999;14:902-7.
  114. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol* 2004;113:59-65.
  115. Kocks JWH, Seys SF, van Duin TS, Diamant Z, Tseligianni IG. Assessing patient-reported outcomes in asthma and COPD patients: which can be recommended in clinical practice? *Curr Opin Pulm Med* 2018;24:18-23.
  116. Nascimento Silva M, Nasipitz C, Sole D. Evaluation of quality of life in children and teenagers with allergic rhinitis: adaptation and validation of the Rhinoconjunctivitis Quality of Life Questionnaire (RQLQ). *Allergol Immunopathol (Madr)* 2001;29:111-8.
  117. Larrosa F, Roura J, Dura MJ, Guirao M, Alberti A, Alobid I. Adaptation and validation of the Spanish version of the Nasal Obstruction Symptom Evaluation (NOSE) scale. *Rhinology* 2015;53:176-80.
  118. van Zijl F, Timman R, Datema FR. Adaptation and validation of the Dutch version of the nasal obstruction symptom evaluation (NOSE) scale. *Eur Arch Otorhinolaryngol* 2017;274:2469-76.
  119. Marro M, Mondina M, Stoll D, de Gabory L. French validation of the NOSE and RhinoQOL questionnaires in the management of nasal obstruction. *Otolaryngol Head Neck Surg* 2011;144:988-93.

120. Karahatay S, Tasli H, Karakoc O, Aydin U, Turker T. Reliability and validity of the Turkish Nose Obstruction Symptom Evaluation (NOSE) scale. *Turk J Med Sci* 2018;48:212-6.
121. Onerci Celebi O, Araz Server E, Yigit O, Longur ES. Adaptation and validation of the Turkish version of the Nasal Obstruction Symptom Evaluation scale. *Int Forum Allergy Rhinol* 2018;8:72-6.
122. Bezerra TF, Padua FG, Pilan RR, Stewart MG, Voegels RL. Cross-cultural adaptation and validation of a quality of life questionnaire: the Nasal Obstruction Symptom Evaluation questionnaire. *Rhinology* 2011;49:227-31.
123. Amer MA, Kabbash IA, Younes A, Elzayat S, Tomoum MO. Validation and cross-cultural adaptation of the Arabic version of the Nasal Obstruction Symptom Evaluation scale. *Laryngoscope* 2017;127:2455-9.
124. Kanatani KT, Slingsby BT, Mukaida K, Kitano H, Adachi Y, Haefner D, et al. Translation and linguistic validation of the Allergy-CONTROL-Score<sup>TM</sup> for use in Japan. *Allergol Int* 2013;62:337-41.
125. Flokstra-de Blok BMJ, Baretta HJ, Fonseca JA, van Heijst E, Kollen BJ, de Kroon J, et al. Control of Allergic Rhinitis and Asthma Test with 1-week recall: validation of paper and electronic version. *Allergy* 2018;73:2381-5.

TABLE E1. Summary of publications on the validation of symptom scores, disease control scores, and QOL scores in AR

| Reference                             | Name and type of score                                                  | Age class of study subjects | Psychometric validity                                                                              |                                                        |                                                                                     |                                                                                                                                     | Responsiveness (ability to detect change) | MCID                                        |          | Other quality indicators                                                                                                                    |
|---------------------------------------|-------------------------------------------------------------------------|-----------------------------|----------------------------------------------------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|---------------------------------------------|----------|---------------------------------------------------------------------------------------------------------------------------------------------|
|                                       |                                                                         |                             | Construct validity (convergent or discriminant, cross-sectional or longitudinal, language)         | Content validity (comprehensiveness of item selection) | Criterion validity (concurrent or predictive vs a criterion standard or an outcome) | Reliability (inter-/intra-rater, test-retest, or internal consistency)                                                              |                                           | Anchor-based                                | SD-based |                                                                                                                                             |
| Wasserfallen et al <sup>55</sup>      | RASS (symptom score)                                                    | Adults                      | ✓                                                                                                  | ✓                                                      | —                                                                                   | ✓<br>Test-retest reliability was 0.8 at 1 wk and internal consistency was 0.8 for individual organs and 0.7 for individual symptoms | ✓                                         | —                                           | —        | Prediction of performance after a change in the number of items; appropriateness of the data distribution and the statistical tests applied |
| Juniper et al <sup>97</sup>           | PRQLQ (QOL score)                                                       | Children                    | ✓                                                                                                  | —                                                      | —                                                                                   | ✓<br>Overall intraclass correlation coefficient for test-retest reliability = 0.93                                                  | ✓                                         | —                                           | —        | —                                                                                                                                           |
| Revicki et al <sup>57</sup>           | RSUI (symptom score)                                                    | Adults                      | ✓                                                                                                  | —                                                      | ✓                                                                                   | —                                                                                                                                   | —                                         | —                                           | —        | —                                                                                                                                           |
| Juniper et al <sup>94</sup>           | RQLQ(S) (QOL score)                                                     | Adults                      | ✓<br>vs diary symptoms, SF-36, “feeling thermometer” patient preference, “standard gamble” utility | —                                                      | —                                                                                   | ✓<br>Intraclass correlation coefficient for test-retest reliability = 0.97                                                          | ✓<br>$P < .001$                           | ✓<br>MID = $0.48 \pm 0.93$ , against a GRCS | —        | —                                                                                                                                           |
| Kemker et al <sup>106</sup>           | AOS (hybrid disease control, symptom, outcome score)                    | Not stated                  | ✓                                                                                                  | —                                                      | —                                                                                   | ✓                                                                                                                                   | ✓                                         | —                                           | —        | —                                                                                                                                           |
| Leong et al <sup>84</sup>             | RQLQ (QOL score) administered to English-speaking patients in Singapore | Adults                      | ✓                                                                                                  | ✓                                                      | —                                                                                   | —                                                                                                                                   | ✓                                         | —                                           | —        | Identification of candidates for scale-shortening                                                                                           |
| Juniper et al <sup>103</sup>          | Mini-RQLQ (QOL score)                                                   | Adults                      | ✓                                                                                                  | ✓                                                      | ✓<br>Against the RQLQ                                                               | ✓                                                                                                                                   | ✓                                         | ✓<br>MID = $0.70 \pm 1.15$ , against a GRCS | —        | —                                                                                                                                           |
| Nascimento Silva et al <sup>116</sup> | Modified RQLQ (QOL score)                                               | Children and teens          | ✓                                                                                                  | ✓                                                      | —                                                                                   | ✓                                                                                                                                   | ✓                                         | —                                           | —        | —                                                                                                                                           |

|                                              |                                                                                                                                                                                                                                                  |                       |   |   |   |                                                                                                                                     |   |                                                                      |   |   |
|----------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|---|---|---|-------------------------------------------------------------------------------------------------------------------------------------|---|----------------------------------------------------------------------|---|---|
| Baiardini et al <sup>95</sup>                | RHINASTHMA (QOL score)                                                                                                                                                                                                                           | Adults                | ✓ | — | ✓ | ✓<br>Test-retest:<br>Pearson<br>coefficient<br>between 0.70 and<br>0.92 for 25 items<br>and between 0.61<br>and 0.69 for 5<br>items | ✓ | —                                                                    | — | — |
| Juniper et al <sup>96</sup>                  | NRQLQ (QOL score)                                                                                                                                                                                                                                | Adults                | ✓ | — | — | ✓                                                                                                                                   | ✓ | —                                                                    | — | — |
| Roberts et al <sup>91</sup>                  | PADQLQ (QOL score)                                                                                                                                                                                                                               | Children<br>and teens | ✓ | — | — | ✓                                                                                                                                   | ✓ | ✓<br>MID = 0.33<br>(95% CI, 0.11<br>to 0.54), against<br>a VAS score | — | — |
| Johnson et al <sup>61</sup>                  | SoAR (symptom score)                                                                                                                                                                                                                             | Adults                | ✓ | — | — | ✓                                                                                                                                   | — | —                                                                    | — | — |
| Stewart et al <sup>50</sup>                  | NOSE (symptom score)<br>in adults (has also<br>been validated in<br>Spanish, <sup>117</sup> Dutch, <sup>118</sup><br>French, <sup>119</sup> Italian,<br>Turkish, <sup>120,121</sup><br>Portuguese, <sup>122</sup> and<br>Arabic <sup>123</sup> ) | Adults                | ✓ | ✓ | ✓ | ✓                                                                                                                                   | ✓ | —                                                                    | — | — |
| Leong et al <sup>85</sup>                    | RQLQ (vs SF-36 in<br>patients with persistent<br>AR)                                                                                                                                                                                             | Adults                | ✓ | — | — | ✓                                                                                                                                   | ✓ | —                                                                    | — | — |
| Okuda et al <sup>86</sup>                    | RQLQ in Japanese and<br>JRQLQ (QOL scores)                                                                                                                                                                                                       | Adults                | ✓ | — | — | ✓                                                                                                                                   | — | —                                                                    | — | — |
| Lo et al <sup>58</sup>                       | RSUI in Chinese patients<br>(symptom score)                                                                                                                                                                                                      | Adults                | ✓ | — | — | ✓                                                                                                                                   | — | —                                                                    | — | — |
| Santanello et al <sup>56</sup>               | NtSS (symptom score)                                                                                                                                                                                                                             | Teens and<br>adults   | ✓ | — | — | ✓                                                                                                                                   | ✓ | —                                                                    | — | — |
| ESPRINT Study<br>Investigators <sup>93</sup> | ESPRINT (QOL score)                                                                                                                                                                                                                              | Adults                | ✓ | ✓ | — | ✓                                                                                                                                   | ✓ | —                                                                    | — | — |
| Stull et al <sup>52</sup>                    | CQ7 (symptom score)                                                                                                                                                                                                                              | Adults                | ✓ | ✓ | — | ✓                                                                                                                                   | ✓ | —                                                                    | — | — |
| AbuRuz et al <sup>90</sup>                   | AAQQ (QOL score)                                                                                                                                                                                                                                 | Adults                | ✓ | ✓ | — | ✓                                                                                                                                   | ✓ | ✓<br>MID ± SD =<br>1.1 ± 1.0, against<br>a GRCS score                | — | — |
| Nogueira-<br>Silva et al <sup>102</sup>      | CARAT, 17 questions<br>(disease control score)                                                                                                                                                                                                   | Adults                | ✓ | ✓ | — | —                                                                                                                                   | — | —                                                                    | — | — |
| Yüksel et al <sup>87</sup>                   | RQLQ and mini-RQLQ<br>in Turkish (QOL<br>scores)                                                                                                                                                                                                 | Adults                | ✓ | — | — | ✓                                                                                                                                   | — | —                                                                    | — | — |
| Yüksel et al <sup>98</sup>                   | PRQLQ in Turkish (QOL<br>score)                                                                                                                                                                                                                  | Children<br>and teens | ✓ | — | — | ✓                                                                                                                                   | — | —                                                                    | — | — |

(continued)

TABLE E1. (Continued)

| Reference                   | Name and type of score                                                                | Age class of study subjects | Psychometric validity                                                                      |                                                        |                                                                                                                                                             |                                                                        |                                           | MCID                                                                                                                                                                   |                                              | Other quality indicators                                               |
|-----------------------------|---------------------------------------------------------------------------------------|-----------------------------|--------------------------------------------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|-------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------|------------------------------------------------------------------------|
|                             |                                                                                       |                             | Construct validity (convergent or discriminant, cross-sectional or longitudinal, language) | Content validity (comprehensiveness of item selection) | Criterion validity (concurrent or predictive vs a criterion standard or an outcome)                                                                         | Reliability (inter-/intra-rater, test-retest, or internal consistency) | Responsiveness (ability to detect change) | Anchor-based                                                                                                                                                           | SD-based                                     |                                                                        |
| Barnes et al <sup>36</sup>  | Mini-RQLQ (QOL score), PNIF (physical measurement), and T4NSS (symptom score)         | Adults                      | —                                                                                          | —                                                      | —                                                                                                                                                           | —                                                                      | —                                         | ✓<br>Direct anchor-based: 0.15 (−0.13 to 0.44); direct anchor-based by regression: 0.23 (0.03 to 0.44); indirect anchor-based by regression: 0.42 units (0.30 to 0.51) | ✓<br>Distribution-based: 0.24 (0.21 to 0.27) | —                                                                      |
| Fonseca et al <sup>67</sup> | CARAT10 to assess rhinitis and asthma in patients with asthma (disease control score) | Adults                      | —                                                                                          | —                                                      | ✓<br>Stated as concurrent validity, vs physician's assessment, VAS of symptoms, and ACQ5 score; probably construct validity, rather than criterion validity | —                                                                      | —                                         | —                                                                                                                                                                      | —                                            | —                                                                      |
| Schatz et al <sup>73</sup>  | RCAT (disease control score)                                                          | Adults                      | ✓                                                                                          | —                                                      | —                                                                                                                                                           | ✓                                                                      | —                                         | —                                                                                                                                                                      | —                                            | —                                                                      |
| Stull et al <sup>53</sup>   | CQ5 (symptom score)                                                                   | Adults                      | ✓                                                                                          | ✓                                                      | —                                                                                                                                                           | ✓                                                                      | —                                         | —                                                                                                                                                                      | —                                            | Determination of the cutoff for detecting congestion                   |
| Demoly et al <sup>78</sup>  | ARCT (disease control score)                                                          | Teens and adults            | ✓                                                                                          | ✓                                                      | —                                                                                                                                                           | ✓                                                                      | ✓                                         | —                                                                                                                                                                      | —                                            | Determination of the cutoff for detecting control Sp, Se, NPV, and PPV |
| Franzke et al <sup>59</sup> | PBI-AR (symptom score)                                                                | Adults                      | ✓                                                                                          | —                                                      | —                                                                                                                                                           | ✓                                                                      | —                                         | —                                                                                                                                                                      | —                                            | Acceptability and feasibility in clinical routine                      |
| Grouin et al <sup>63</sup>  | AdSS (symptom score)                                                                  | Children, teens, and adults | ✓<br>✓<br>vsRTSS and a combined score                                                      | —                                                      | —                                                                                                                                                           | —                                                                      | —                                         | —                                                                                                                                                                      | —                                            | —                                                                      |

|                                 |                                                                   |                    |   |   |   |   |   |                                       |                                 |                                 |
|---------------------------------|-------------------------------------------------------------------|--------------------|---|---|---|---|---|---------------------------------------|---------------------------------|---------------------------------|
| Hafner et al <sup>65</sup>      | ACS (a CSMS)                                                      | Adults             | ✓ | — | — | ✓ | — | —                                     | —                               | Feasibility                     |
| Kiotseridis et al <sup>92</sup> | PADQLQ (QOL score)                                                | Children and teens | ✓ | — | — | ✓ | ✓ | —                                     | ✓<br>0.5<br>SD = 0.42           | —                               |
| Valero et al <sup>54</sup>      | CQ7 (in Spanish) (symptom score)                                  | Adults             | ✓ | — | — | ✓ | ✓ | —                                     | —                               | Sp, Se, cutoff                  |
| Braido et al <sup>89</sup>      | RAPP (QOL score)                                                  | Teens and adults   | ✓ | ✓ | ✓ | ✓ | ✓ | —                                     | ✓<br>Cutoff of 2                | —                               |
| Fonseca et al <sup>68</sup>     | CARAT10 (disease control score)                                   | Adults             | ✓ | — | — | ✓ | ✓ | —                                     | —                               | —                               |
| Hafner et al <sup>66</sup>      | ACS for the eyes and nose only (RC-ACS, a CSMS)                   | Adults             | ✓ | — | — | ✓ | — | —                                     | —                               | —                               |
| Schatz et al <sup>74</sup>      | Mini-RQLQ (QOL score)                                             | Adults             | — | ✓ | — | ✓ | — | —                                     | —                               | —                               |
|                                 | RCAT (disease control score)                                      | Adults             | — | ✓ | — | ✓ | — | ✓<br>2.6 for the RCAT                 | —                               | —                               |
| Demoly et al <sup>82</sup>      | VAS for the burden of AR in primary care (disease control score)  | Adults             | ✓ | — | — | — | ✓ | —                                     | —                               | —                               |
| Higaki et al <sup>64</sup>      | A T5SS (in Japanese cedar/cypress pollinosis) (symptom score)     | Adults             | — | — | — | — | — | ✓<br>1.426 in 2009 and 1.441 in 2010  | —                               | —                               |
|                                 | A T6SS (in Japanese cedar/cypress pollinosis) (symptom score)     | Adults             | — | — | — | — | — | ✓<br>4.115 in 2009 and 3.183 in 2010  | —                               | —                               |
|                                 | JRQLQ, 17-item (in Japanese cedar/cypress pollinosis) (QOL score) | Adults             | — | — | — | — | — | ✓<br>10.469 in 2009 and 6.026 in 2010 | —                               | —                               |
| Kanatani et al <sup>124</sup>   | ACS in Japanese (a CSMS)                                          | Teens and adults   | — | ✓ | — | — | — | —                                     | —                               | Accessibility and acceptability |
| Meltzer et al <sup>75</sup>     | RCAT in rhinitis (disease control score)                          | Teens and adults   | ✓ | — | — | ✓ | — | ✓<br>Mean of 14 measurements: 2.4     | ✓<br>SEM = 2.1;<br>0.5 SD = 2.2 | Screening accuracy              |
| Borrego et al <sup>70</sup>     | CARATKids (disease control score)                                 | Children           | — | ✓ | — | — | — | —                                     | —                               | Cognitive testing               |

(continued)

TABLE E1. (Continued)

| Reference                         | Name and type of score                                    | Age class of study subjects | Psychometric validity                                                                      |                                                        |                                                                                     |                                                                        |                                           | MCID                                                                                                                                                    |                                                              | Other quality indicators    |
|-----------------------------------|-----------------------------------------------------------|-----------------------------|--------------------------------------------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------------------------------------|------------------------------------------------------------------------|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|-----------------------------|
|                                   |                                                           |                             | Construct validity (convergent or discriminant, cross-sectional or longitudinal, language) | Content validity (comprehensiveness of item selection) | Criterion validity (concurrent or predictive vs a criterion standard or an outcome) | Reliability (inter-/intra-rater, test-retest, or internal consistency) | Responsiveness (ability to detect change) | Anchor-based                                                                                                                                            | SD-based                                                     |                             |
| Devillier et al <sup>62</sup>     | RTSS (symptom score)                                      | Children, teens, and adults | —                                                                                          | —                                                      | —                                                                                   | —                                                                      | —                                         | ✓<br>1.24 ± 0.17 (vs GRCS) and 1.12 ± 0.14 (vs RQLQ) in children, 1.33 ± 0.14 and 1.20 ± 0.13 in adolescents; and 1.13 ± 0.14 and 0.89 ± 0.12 in adults | ✓<br>0.33 SD = from 1.09 to 1.13; 0.5 SD = from 1.22 to 1.40 | —                           |
| Linhares et al <sup>47</sup>      | CARATKids (disease control score)                         | Children                    | ✓                                                                                          | ✓                                                      | —                                                                                   | ✓                                                                      | ✓                                         | —                                                                                                                                                       | —                                                            | —                           |
| Wang et al <sup>80</sup>          | ARCT in Chinese (disease control score)                   | Adults                      | ✓                                                                                          | —                                                      | —                                                                                   | —                                                                      | —                                         | —                                                                                                                                                       | —                                                            | —                           |
| Demoly et al <sup>60</sup>        | PBI-AR (symptom score)                                    | Adults                      | ✓                                                                                          | —                                                      | ✓<br>Against a discomfort VAS and the mini-RQLQ                                     | ✓                                                                      | —                                         | —                                                                                                                                                       | ✓<br>0.5 SD = 0.4                                            | —                           |
| van der Leeuw et al <sup>69</sup> | CARAT10 in Dutch (disease control score)                  | Adults                      | ✓                                                                                          | —                                                      | —                                                                                   | ✓                                                                      | —                                         | ✓<br>3.50 against the GRCS                                                                                                                              | ✓<br>SEM = 2.83                                              | —                           |
| Zicari et al <sup>51</sup>        | NOSE in children (symptom score)                          | Children                    | ✓                                                                                          | —                                                      | —                                                                                   | —                                                                      | —                                         | —                                                                                                                                                       | —                                                            | —                           |
| Boulay and Boulet <sup>77</sup>   | RCSS (disease control score)                              | Adults                      | ✓<br>Against RQLQ(S), VAS, and TNSS-TRSS                                                   | ✓                                                      | ✓                                                                                   | ✓                                                                      | —                                         | —                                                                                                                                                       | —                                                            | —                           |
| Fernandes et al <sup>76</sup>     | RCAT in Brazilian Portuguese (disease control score)      | Teens                       | ✓                                                                                          | —                                                      | —                                                                                   | ✓                                                                      | —                                         | —                                                                                                                                                       | —                                                            | Sensitivity and specificity |
| Mavroudi et al <sup>99</sup>      | Ped-AR-QOL (QOL score)                                    | Children and teens          | ✓<br>With the Disabkids pediatric QOL scale                                                | —                                                      | —                                                                                   | ✓                                                                      | —                                         | —                                                                                                                                                       | —                                                            | —                           |
| Wang et al <sup>79</sup>          | ARCT in Chinese (disease control score)                   | Children, teens, and adults | —                                                                                          | —                                                      | —                                                                                   | —                                                                      | ✓                                         | —                                                                                                                                                       | —                                                            | —                           |
| Amaral et al <sup>71</sup>        | CARATKids in Brazilian Portuguese (disease control score) | Children                    | ✓                                                                                          | —                                                      | —                                                                                   | ✓                                                                      | ✓                                         | —                                                                                                                                                       | —                                                            | Sensitivity and specificity |

|                                       |                                                                                                    |                    |   |   |   |   |   |   |                          |                               |
|---------------------------------------|----------------------------------------------------------------------------------------------------|--------------------|---|---|---|---|---|---|--------------------------|-------------------------------|
| Caimmi et al <sup>83</sup>            | Four VASs measuring control of overall, nasal, ocular, and asthma symptoms (disease control score) | Not stated         | — | — | — | ✓ | — | — | —                        | Sensitivity and acceptability |
| Emons et al <sup>72</sup>             | CARATKids in Dutch (disease control score)                                                         | Children and teens | ✓ | — | — | ✓ | ✓ | ✓ | GRCS: 2.76;<br>ACT: 2.77 | 1.96<br>SEM = 2.80            |
| Molinengo et al <sup>88</sup>         | RAPP (QOL score)                                                                                   | Adults             | ✓ | ✓ | — | — | — | — | —                        | —                             |
| Liedtke et al <sup>101</sup>          | RCAT (disease control score)                                                                       | Adults             | ✓ | — | — | — | — | — | —                        | —                             |
| Zhu et al <sup>81</sup>               | ARCT in Chinese (disease control score)                                                            | Adults             | ✓ | — | — | — | ✓ | — | —                        | —                             |
| Flokstra-de Blok et al <sup>125</sup> | CARAT paper and electronic versions (disease control score)                                        | Adults             | ✓ | — | — | ✓ | — | — | —                        | —                             |

*AAQO*, Arabic Allergic Rhinitis Quality of Life Questionnaire; *AdSS*, adjusted symptom score; *ESPRINT*, Cuestionario ESPañol de Calidad de Vida en RINiTis; *GRCS*, global rating of change scale; *JRQLQ*, Japan Rhinoconjunctivitis Quality of Life Questionnaire; *MID*, minimal important difference; *NPV*, negative predictive value; *NRQLQ*, Nocturnal Rhinoconjunctivitis Quality of Life Questionnaire; *NtSS*, nighttime symptoms score; *PADQLQ*, Pediatric Allergic Disease Quality of Life Questionnaire; *PBI-AR*, Patient Benefit Index—Allergic Rhinitis; *Ped-AR-QOL*, Pediatric Allergic Rhinitis Quality of Life Questionnaire; *PNIF*, peak nasal inspiratory flow; *PPV*, positive predictive value; *PRQLQ*, Paediatric Rhinoconjunctivitis Quality of Life Questionnaire; *RAPP*, RhinAsthma Patient Perspective; *RASS*, rhinoconjunctivitis and asthma symptom score; *RHINASTHMA*, a disease-specific QOL questionnaire; *RQLQ*, Rhinoconjunctivitis Quality of Life Questionnaire; *RQLQ(S)*, Standardized Rhinoconjunctivitis Quality of Life Questionnaire; *RSUI*, Rhinitis Symptom Utility Index; *SF-36*, short form 36 health survey; *SoAR*, Scale of Allergic Rhinitis; *T4NSS*, a 4-item total nasal symptom score.

*Note.* Age classes of study subjects are defined as follows: children, younger than 13 y; teens, aged 13-17 y; adults, aged 18 y and older.

Reproduced with permission of copyright owner. Further reproduction  
prohibited without permission.