

Adding measures to the asthma outcome measurement toolbox: New findings, new issues

Sandra R. Wilson, PhD *Palo Alto, Calif*

Key words: Composite measures, asthma outcomes, asthma control, asthma symptoms, test-retest reliability

The report of the 2010 Asthma Outcomes Workshop convened by the National Institutes of Health and the Agency for Healthcare Research and Quality was published earlier this year in a *Journal of Allergy and Clinical Immunology* supplement.¹ The workshop's premise, like that of 2 previous efforts,^{2,3} was that "clinical research will highly benefit from standardization of major outcomes in terms of definition and assessment methodology" because this "will permit useful comparisons across interventional or observational studies and will allow more effective data sharing."

Workshop recommendations were made in 7 areas: biomarkers, pulmonary physiology, composite scores of asthma control, symptoms, exacerbations, health care use and costs, and disease effect on quality of life. In each area the report identified the currently available measures and measurement procedures for which there was published psychometric/performance data; evaluated the evidence as to their reliability, their validity for their intended purpose, and the adequacy of their other psychometric properties; identified information that was not available regarding their performance characteristics; identified measures that qualified as core (to be required) or supplementary measures for federally sponsored asthma clinical research; and determined whether new, better measurement instruments are needed. Two articles in this issue of the *Journal of Allergy and Clinical Immunology* speak to questions raised in the Workshop report's "Symptoms" and "Composite measures" chapters.^{4,5}

The "Symptoms" report concluded that evaluation of symptoms is a core measurement that should be made in both experimental and observational asthma clinical research.⁶ However, it did not recommend any specific measurement instrument or procedure. The Asthma Symptom Utility Index (ASUI), the most widely used retrospective symptom recall questionnaire for adults, was recommended as a supplementary measure for optional use depending on the aims of a particular research study. The ASUI was seen to have good reported measurement characteristics⁷ but lacked information on fundamental psychometric properties. In this issue Bime et al⁴ have provided substantial new information on the measurement properties of the ASUI

using data from the Study of Acid Reflux and Asthma (SARA) and Safety of Inactivated Influenza Vaccine in Asthma (SIIVA) trials, including the ASUI's internal consistency reliability (coefficient α), test-retest reliability (intraclass correlation coefficient), construct validity (the correlations between the ASUI and both the Asthma Control Questionnaire [ACQ] and mini Asthma Quality of Life Questionnaire by using baseline scores of SARA participants), predictive validity (the ability of baseline ASUI scores to predict the frequency of episodes of poor asthma control within the subsequent 2 weeks), responsiveness to change in FEV₁ and ACQ scores of SARA trial participants assessed over consecutive study visits, and minimal clinically important score difference estimated by using anchor-based and distribution-based analyses.

The properties of the ASUI that these analyses reveal should increase confidence in the use of this instrument in asthma clinical research, although the method of assessing test-retest reliability using only the subset of patients who remained “stable” in terms of other concurrently assessed outcome measures, an unfortunately common practice, tends to overestimate the true value of the intraclass correlation coefficient.⁸ Importantly, the ASUI is not encumbered by copyright restrictions, and although scoring is more complicated than a simple summation of responses, with appropriate computer support, this should be a minor consideration. The new ASUI performance results also come from US study samples with a higher proportion of persons of ethnic minority background and lower educational level than was the case in the sample from which the original preference and utility data and psychometric results were obtained.

Flood et al⁹ found that preference weights, utility functions, and mean ASUI scores for European samples differed from those of the US developmental sample. Although this might appear to indicate some limitation of the ASUI scoring when it comes to international use, in fact, the samples all were very small (n = 30 each) and not selected to be representative of their respective countries. Sex, age, education, asthma severity, and comorbidities differed substantially among the samples of English, French, and Italian patients and differed from the US sample, and the preferences and ASUI means were not adjusted for these characteristics. Thus the evidence of country-related differences is much weaker than it might appear. As Bime et al⁴ note, in both the US and European samples, ASUI scores were monotonically associated with asthma severity and had comparable levels of association with other asthma outcome measures. This finding is the most telling because the consistency of these associations is clearly a more important consideration in terms of the wider use of the ASUI than are the sample differences in observed preference and mean ASUI scores. Retaining the ability to compare study populations on the same measurement instrument and scale might well outweigh any theoretic advantage of scoring based on locally relevant preference weights and utilities. It is worth noting that the

From the Department of Health Services Research, Palo Alto Medical Foundation Research Institute.

Disclosure of potential conflict of interest: S. R. Wilson declares that she has no relevant conflicts of interest.

Received for publication August 31, 2012; accepted for publication September 4, 2012.

Corresponding author: Sandra R. Wilson, PhD, Department of Health Services Research, Palo Alto Medical Foundation Research Institute, 795 El Camino Real (Ames Bldg), Palo Alto, CA 94301. E-mail: wilsons@pamfri.org.

J Allergy Clin Immunol 2012;■■■:■■■-■■■.

0091-6749/\$36.00

© 2012 American Academy of Allergy, Asthma & Immunology

<http://dx.doi.org/10.1016/j.jaci.2012.09.002>

use of any utility function in scoring inherently ignores individual differences in preferences and utilities to gain the advantages associated with using average population preference values.

A new measure, the Asthma Disease Activity Score (ADAS), is reported in the article by Greenberg et al⁵ and was not available for review in the workshop's "Composite measures" report.¹⁰ Like all composite measures, the ADAS-6 combines information on multiple constructs typically obtained from multiple measurement procedures. The Asthma Control Questionnaire (ACQ), for example, requires spirometry to obtain FEV₁ values and a brief questionnaire to obtain information on multiple indicators of asthma control.¹¹ Combining reliable measures of relevant constructs can contribute to a measure that has greater predictive value with regard to other outcomes than any of its constituents alone, thus making an additional contribution to the outcome measurement toolbox. In fact, the weighting of the constituents of such composite scores is often derived from multivariate predictive modeling of the outcome of interest. The 6-item version (ADAS-6) provides a continuous score derived from measures of rescue β -agonist use (mean self-reported puffs per day from daily diary data), percent predicted FEV₁ (from office spirometry), symptoms (from the symptom domain score of the Juniper Asthma Quality of Life Questionnaire, or AQLQ),¹¹ reported nocturnal awakening (from daily diary data), and 2 variables: diurnal variability in rescue β -agonist use and diurnal variability in peak expiratory flow rate (from diurnal diary data). The latter 2 variables are omitted in the 4-item score (ADAS-4). The ADAS-6 has reasonable internal consistency reliability, is responsive to treatment, distinguishes well between asthma of different levels of severity (as does the ADAS-4), and is predictive of future exacerbations.

Concern about a potentially increased subject burden with the use of composite measures is lessened if the measurements that go into the composite are individually required for asthma research (eg, symptoms and FEV₁). The ADAS requires data from multiple sources, including one subscore of the AQLQ, which has use restrictions that require administration of the entire questionnaire.¹² Such a composite instrument faces a higher burden of proof of its value. Respondent burden is only partially addressed by testing a shortened (ADAS-4) version that avoids the need for patients to keep diurnal diaries. However, the ADAS could potentially use an alternative symptom measure, such as

the ASUI. The ADAS was designed to measure disease activity, and although the authors caution that this is not the same construct as asthma control, the information going into the ADAS overlaps that of the Asthma Control Test and ACQ, both of which were recommended as core measures. Hence it will be important to determine whether the ADAS, with its additional use of diary data and differential weighting of its components, provides unique predictive information not provided by the Asthma Control Test and ACQ. Greenberg et al⁵ have provided very useful information and insights that might lead to more effective composite measures and more efficient measurement of core asthma outcomes.

REFERENCES

1. Busse WW, Morgan WJ, Taggart V, Togias A. Asthma outcomes workshop: overview. *J Allergy Clin Immunol* 2012;129(suppl):S1-8.
2. Bailey WC, Wilson SR, Weiss KB, Windsor RA, Wolle JM. Measures for use in asthma clinical research: overview of the NIH workshop. *Am J Respir Crit Care Med* 1994;149(suppl):S1-8.
3. Reddel HK, Taylor DR, Bateman ED, Boulet LP, Boushey HA, Busse WW, et al. An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med* 2009;180:59-99.
4. Bime C, Wei CY, Holbrook KT, Sockrider MM, Revicki DA, Wise RA. Asthma Symptom Utility Index: reliability, validity, responsiveness and the minimal important difference in adult asthma patients. *J Allergy Clin Immunol* 2012. ■■■■■.
5. Greenberg S, Liu N, Kaur A, Lakshminarayanan M, Zhou Y, Nelsen L, et al. The Asthma Disease Activity Score: a discriminating, responsive measure predicts future asthma attacks. *J Allergy Clin Immunol* 2012. ■■■■■.
6. Krishnan JA, Lemanske RF, Canino GJ, Elward KS, Kattan M, Matsur EC, et al. Asthma outcomes: symptoms. *J Allergy Clin Immunol* 2012;129(suppl):S124-35.
7. Revicki DA, Leidy NK, Brennan-Diener F, Sorensen S, Togias A. Integrating patient preferences into health outcomes assessment: the multiattribute Asthma Symptom Utility Index. *Chest* 1998;114:998-1007.
8. Norman G, Stratford P, Regeher G. Biases in the retrospective calculation of reliability and responsiveness from longitudinal studies. In: Mesbah M, Cole B, Ting Lee M-L, editors. *Statistical methods for quality of life studies: design, measurements, and analysis*. Boston: Kluwer Academic Publishers; 2002. p. 21-31.
9. Flood EM, De Cock E, Mörk A-C, Revicki DA. Evaluating preference weights for the Asthma Symptom utility Index (ASUI) across countries. *Health Qual Life Outcomes* 2006;4:51.
10. Cloutier MM, Schatz M, Castro M, Clark N, Kelly HW, Mangione-Smith R, et al. Asthma outcomes: composite scores of asthma control. *J Allergy Clin Immunol* 2012;129(suppl):S24-33.
11. 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.
12. Juniper EF, Buist AS, Cox FM, Ferrie PJ, King DR. Validation of a standardized version of the Asthma Quality of Life Questionnaire. *Chest* 1999;115:1265-70.