

Psychometric Properties of the Asthma Symptom Diary (ASD), a Diary for Use in Clinical Trials of Persistent Asthma



Gary Globe, PhD, MBA^a, Ingela Wiklund, PhD^b, Joseph Lin, MD^c, Wen-Hung Chen, PhD^d, Mona Martin, RN, MPA^e, Maria S. Mattera, MPH^d, Robyn von Maltzahn, MSc^b, Jing Yuan Feng, MS^a, Yun Chon, PhD^a, Hema N. Viswanathan, PhD^a, and Michael Schatz, MD, MS^f *Thousand Oaks, Foster City, San Diego, Calif; London, United Kingdom; Bethesda, Md; and Mountlake Terrace, Wash*

What is already known about this topic? Asthma symptoms are highly variable, and symptom burden does not correlate well with clinical parameters. The asthma symptom diary was created with well-documented patient input.

What does this article add to our knowledge? This article documents the psychometric validation of a new valid and reliable patient-reported outcome asthma symptom measure.

How does this study impact current management guidelines? This study provides an asthma symptom measure that has well-documented content validity and good psychometric properties.

BACKGROUND: No currently available asthma symptom diary has sufficient validation to be recommended for use as a core asthma outcome measure.

OBJECTIVE: The objective of this study was to provide validation data for the 10-item asthma symptom diary (ASD).

METHODS: Data were collected in a 4-week prospective, observational study. Subjects completed 3 study visits, completing the ASD twice daily at home for 28 days. Psychometric properties in terms of dimensionality, reliability, validity, and responsiveness were assessed.

RESULTS: Data from 276 subjects were analyzed; mean age was 42.9 (standard deviation [SD] = 16.4) years, mean asthma duration was 23.3 (SD = 16.8) years, and 69.6% were female. Confirmatory factor and Rasch analysis supported the ASD as unidimensional and adequately measuring the spectrum of asthma symptom severity. High Cronbach's α (0.94) and intraclass correlation coefficients (0.89-0.95) supported reliability. A high correlation between the 7-day average ASD score and the Asthma Control Questionnaire (ACQ) total score ($r = 0.75$) and Asthma Quality of Life Questionnaire total scores ($r = -0.76$), and a moderate correlation with FEV₁% predicted ($r = -0.30$) supported convergent validity. Significant differences ($P < .001$) between groups classified by ACQ scores supported known-group validity. The 7-day average ASD scores were responsive to change, with significantly higher score changes ($P < .001$) in responders versus nonresponders. Minimally important differences were calculated and found to be in the range of 0.1-0.3.

CONCLUSION: Results of this study indicated that the ASD is a reliable and valid asthma symptom measure for use in adult and adolescent asthma patients to evaluate the effect of treatment on asthma in clinical trials. © 2015 American Academy of Allergy, Asthma & Immunology (J Allergy Clin Immunol Pract 2016;4:60-6)

Key words: Asthma; Symptoms; Patient-reported outcome; Psychometrics; Validity

Asthma is a chronic respiratory illness characterized by airway remodeling, inflammation, and immunological hyper-responsiveness to allergens.¹ Approximately 300 million people worldwide suffer from this condition.² According to data from the 2011 National Health Interview Survey, asthma affected 17.5 million adults and 7.1 million children in the United States in 2009—a prevalence of 8.2% that has increased 1.2% annually

^aAmgen Inc., Thousand Oaks, Calif

^bEvidera, London, United Kingdom

^cGilead Sciences, Inc., Foster City, Calif

^dEvidera, Bethesda, Md

^eHealth Research Associates, Inc., Mountlake Terrace, Wash

^fDepartment of Allergy, Kaiser Permanente Medical Center, San Diego, Calif

The study was funded by Amgen Inc.

Conflicts of interest: G. Globe is employed by and has stock/stock options in Amgen.

I. Wiklund has received payment for manuscript preparation from Evidera. J. Lin was employed by and has stock/stock options in Amgen. W.-H. Chen has received travel support and was employed by Evidera. M. Martin has a contract to conduct research with Amgen. M. S. Mattera is employed by Evidera. R. von Maltzahn has received fees from Amgen as Amgen contracted with Evidera for the authors to assist in writing the manuscript; Amgen paid Evidera to perform this work. J. Y. Feng was employed by Amgen. Y. Chon is employed by Amgen. H. N. Viswanathan is employed by and has stocks in Amgen. M. Schatz has received consultancy fees and travel support from Amgen.

Received for publication May 16, 2014; revised July 16, 2015; accepted for publication July 22, 2015.

Available online November 6, 2015.

Corresponding author: Gary Globe, PhD, MBA, Global Health Economics, Amgen Inc., One Amgen Center Drive MS 28-3, Thousand Oaks, CA 91320. E-mail: gglobe@amgen.com.

2213-2198

© 2015 American Academy of Allergy, Asthma & Immunology

<http://dx.doi.org/10.1016/j.jaip.2015.07.008>

Abbreviations used

ACQ- Asthma Control Questionnaire
AQLQ- Asthma Quality of Life Questionnaire
ASD- Asthma symptom diary
CDF- Cumulative distribution function
CFA- Confirmatory factor analysis
CFI- Comparative fit index
DIF- Differential item functioning
FEV₁- Forced expiratory volume in 1 second
ICC- Intraclass correlation coefficient
MID- Minimum important difference
PGIC- Patient global impression of change
PRO- Patient-reported outcome
RMSEA- Root mean square error of approximation
SD- Standard deviation
SEM- Standard error of measurement
WPAI- Work Productivity and Activity Impairment Questionnaire
WRMR- Weighted root mean square residual

between 2001 and 2009.³ Uncontrolled asthmatics have been found to experience significantly reduced quality of life, as well as work productivity.⁴ As the burden experienced by a patient as a result of a symptom does not correlate well with clinical measures,⁵ direct patient reports, captured through patient-reported outcome (PRO) measures, are integral to defining asthma status and risk.^{1,4} Asthma symptom diaries are recommended for research purposes, specifically for use in clinical trials to capture the variable nature of asthma symptoms. However, no currently available asthma symptom diary has sufficient evidence of validity to be recommended for use as a core asthma outcome measure.⁶ A reliable and valid daily symptom diary may also be useful in clinical practice for patients with particularly severe or labile disease as it captures the variable nature of the disease and the patients' daily state on treatment.¹

The aim of this study was to assess the validation properties of the Amgen asthma symptom diary (ASD). The original development of the ASD, involving qualitative concept elicitation and cognitive interviews with patients with asthma, is described in detail elsewhere.⁷

METHODS

Study design

This study was a 4-week multicenter, prospective, observational study, in which treatment was provided at the discretion of the individual investigator. A total of 276 adult (≥ 18 years) and adolescent subjects (≥ 12 years) with clinically diagnosed persistent asthma were recruited. The study objective was to assess the psychometric properties of the ASD. At each study visit (visit 1 [baseline]: day 1; visit 2: day 14 $\pm$ 3 days [ie, day 11-day 17]; visit 3: day 28 $\pm$ 3 days [ie, day 25-day 31]), subjects completed the Asthma Control Questionnaire (ACQ),⁸ the Asthma Quality of Life Questionnaire (AQLQ12+),⁹ and the Work Productivity and Activity Impairment Questionnaire: Asthma (WPAI-Asthma).¹⁰ At all 3 study visits, clinical site principal investigators or designees performed spirometry after the participant completed the ACQ. At visit 1, a baseline clinical form was completed that captured clinical variables as well as the clinician's global impression of control (CGIC). At visits 2 and 3, subjects completed the patient global impression of change (PGIC) to assess severity of asthma symptoms since the last study visit. A follow-up clinical form was completed at visits 2 and 3, and

this captured adherence to the ASD, exacerbations, change in controller medication, and a clinician assessment of change in asthma control (CAC) since the previous visit (Figure E1, available in this article's Online Repository at www.jaci-inpractice.org). During the study, subjects completed the ASD at home on an electronic device twice daily (upon awakening before taking asthma medication and in the evening before retiring to bed), beginning on the evening of visit 1 (day 1) and continuing until visit 3 (day 28 $\pm$ 3 days). Further clinical visit details are available in Table E1 in this article's Online Repository at www.jaci-inpractice.org. Institutional Review Board approval was obtained before study initiation.

Description of the asthma symptom diary

The ASD was developed based on concept elicitation and cognitive interviews with patients with clinically diagnosed persistent asthma. Five morning questions (denoted as AM items) assess nighttime symptom severity in relation to wheezing, shortness of breath, cough, and chest tightness, and the frequency of nighttime awakening. An additional morning item assessed duration of nighttime awakening (Table 4 in Globe et al⁷). However, after an initial analysis of the data, this item was found to have considerable overlap with the frequency of nighttime awakening item. This finding was confirmed in the original qualitative patient interviews, and through consultation with clinical experts, the item was removed to leave 5 morning items. A further 5 evening questions (denoted as PM items) assess symptom severity in relation to wheezing, shortness of breath, cough, and chest tightness, and activity limitation since waking. Items are scored from "0" (no symptom, no nighttime awakening, or no activity limitation) to "4" (very severe symptom, unable to sleep, or extreme activity limitation). A daily ASD score is the mean of the 10 items, and the mean of 7 consecutive daily scores is calculated as the "7-day average ASD score." Responses for all 10 items are required to calculate the daily ASD score; otherwise, it is treated as missing. Compliance was captured electronically. For the 7-day average asthma symptom score, scoring is done with no imputation using the mean of at least 4 of the 7 daily ASD scores as a mean weekly item score. The 7-day average ASD score ranges from 0 to 4.

Patient global rating of change. The PGIC is a single-item measure of the perceived change in asthma course since the prior visit with 7 response categories from "extremely better," to "no change," to "extremely worse."^{11,12}

Clinician global rating of change. The CGIC is a single-item measure of the clinician-rated change in the patient's asthma since the prior visit with 5 response categories: "much worse," "worse," "no change," "better," and "much better."¹¹

Statistical analyses

Classical test theory¹³ as well as modern psychometric theory,¹⁴ including Rasch analysis, was used to examine the psychometric properties of the 10-item ASD. Table 1 describes the statistical analysis used.

Item performance. The performance of the 10 ASD items was evaluated by means of descriptive item statistics (mean, standard deviation [SD]), percentage of lowest and highest responses (floor and ceiling effects, respectively), and item frequency distribution at visit 1.

Confirmatory factor analysis. Confirmatory factor analysis (CFA)¹⁵ was conducted using visit 1 data to assess whether the ASD items fit the hypothesized unidimensional construct using Mplus software (Muthén & Muthén, Los Angeles, Calif).¹⁶ The CFA

TABLE I. Summary of statistical analysis

Statistical analysis	Description
Confirmatory factor analysis	Tests whether the data fit a hypothesized measurement model by examining the variance-covariance matrix of correlated variables. Items of the same scale should all belong to a unidimensional model
Rasch model analysis	Use of mathematical models to describe items in terms of item difficulty and discrimination parameters and to transform the categorical item responses into an interval score scale. The item parameters allow us to examine whether the items display necessary psychometric properties
Differential item functioning	Refers to when people from different groups (eg, gender or ethnicity) with the same latent trait (ability and/or skill) have different probabilities of selecting the same item response, sometimes also referred to as measurement bias
Reliability (internal consistency using Cronbach's α)	Internal consistency reliability explores the degree to which the items are associated with each other. Items assessing the same concept should show high internal consistency
Reliability (test-retest reliability)	Stability of scores over time when no change is expected in the concept of interest. A measure is said to have high test-retest reliability when the scores are stable over time
Concurrent validity	Strength of correlation between the study measure and other measures of the same concept. The validity of the study measure is supported when its score is highly correlated with other similar measures
Known-groups validity	Known-groups validity assesses the degree to which the measure can distinguish among groups of subjects hypothesized to be different in the concept of interest, such as level of disease severity
Responsiveness	Evidence that a measure can identify differences in scores over time in individuals or groups who have changed with respect to the measurement concept
Minimally important difference	Estimate of the smallest change in a treatment outcome that a patient would identify as important
Responder definition	Estimate of the individual patient score change over a predetermined time period that should be interpreted as a meaningful improvement

model fit was assessed with the comparative fit index (CFI > 0.90), root mean square error of approximation (RMSEA < 0.05), and weighted root mean square residual (WRMR close to 1).

Rasch model analysis. Rasch model analysis, using RUMM2030,¹⁷ was conducted using visit 1 data to evaluate whether the ASD items fit to the hypothesized unidimensional measurement model. The overall model fit was determined by the overall χ^2 fit statistic. The item fit was evaluated through nonsignificant item-level χ^2 fit statistics (following adjustment for multiple comparisons) and item fit residuals between -3.0 and 3.0. The analysis was also used to assess whether the item response categories were appropriately ordered, as indicated by nonoverlapping category threshold parameters.

Differential item functioning. Differential item functioning (DIF) analysis using ordinal logistic regression¹⁸ was conducted using SAS (SAS Institute, Cary, NC) to evaluate whether the items in the ASD performed differently between male and female subjects while controlling for total ASD scores. DIF by sex for an item was indicated by a significant χ^2 statistic (following adjustment for multiple comparisons).

Reliability. Cronbach's coefficient α ¹⁹ and a summary of α levels with each constituent item deleted were used to report the internal consistency reliability for the ASD. Reproducibility of the ASD was examined assessing test-retest reliability by an intraclass correlation coefficient (ICC) within stable subjects (defined by "no change" on the PGIC; FEV₁ absolute percent change < 8%; and ACQ change < 0.5).

Construct validity. We tested construct validity specifically by evaluating convergent validity and known-groups validity, as recommended in the Food and Drug Administration (FDA) Guidance document.²⁰ Convergent validity of the ASD was assessed at baseline using a Spearman correlation with ACQ, AQLQ12+ (overall and domain scores), WPAI-Asthma (overall work impairment and activity impairment), and FEV₁% predicted. Known-groups validity

was examined by the analysis of covariance comparing the mean of the 7-day average ASD scores at baseline by clinician global impression of asthma control ("well-controlled," "not well-controlled," and "very poorly controlled" asthma), by the ACQ score (≤ 0.75 , between >0.75 and <1.5 , and ≥ 1.5), and by the AQLQ12+ score (above and below the median).

Responsiveness. Responsiveness is a key criterion for measures used in clinical trials. Given the observational study design, a change was evaluated in subsets of patients in whom predefined changes occurred. The ASD's ability to detect a change was assessed by comparing changes in 7-day average ASD scores from visit 1 to visit 3 between responders and nonresponders. Responders versus nonresponders were defined in 3 ways: clinician rating of CAC (responders defined as subjects with clinician rating of asthma control as "better" and "much better" at visit 3; nonresponders defined as subjects with clinician ratings of asthma control as "no change," "worse," or "much worse" at visit 3); PGIC ratings (responders defined as subjects with PGIC ratings of "extremely better," "moderately better," or "a little better" at visit 3; nonresponders defined as subjects with all other ratings, eg, "no change" to "extremely worse"); and a change in the ACQ score (responders defined as subjects with a decrease in the ACQ score ≥ 0.5 from visit 1 to visit 3; nonresponders defined as subjects with an increase in ACQ score or a decrease in the ACQ score < 0.5).

Exploratory analysis of minimum important difference and responder definition. Both anchor-based and distribution-based methods were used to explore a preliminary minimum important difference (MID). For anchor-based methods, 3 external anchors for MID from visit 1 to visit 2 and from visit 2 to visit 3 were used: (1) subjects rated as "better" controlled by their clinician; (2) subjects who self-rated themselves as "a little better" by the PGIC; and (3) subjects with an ACQ change score ≥ 0.5 but < 0.75.^{21,22} Two distribution-based methods were used: (1) a 0.5 SD and (2) 1 standard error of measurement (SEM) of the 7-day average ASD score at visit 1.²³ The percentage of subjects meeting the various MID definitions was calculated.

TABLE II. Subject demographic and clinical characteristics at visit 1 (N = 276)

Characteristics	Visit 1
Age, y	
Mean (SD)	42.9 (16.4)
Median (min, max)	46.0 (12.0-70.0)
(Q1, Q3)	(29.5, 57.5)
Adolescent (aged 12-17), n (%)	26 (9.4%)
Adult (aged 18+), n (%)	250 (90.6%)
Gender n (%)	
Female	192 (69.6%)
Ethnicity, n (%)	
Hispanic or Latino	30 (10.9%)
Not Hispanic or Latino	246 (89.1%)
Race, n (%)	
African American and/or African Heritage	59 (21.4%)
American Indian or Alaskan Native	7 (2.5%)
Asian	4 (1.4%)
Native Hawaiian or other Pacific Islander	2 (0.7%)
White or Caucasian	195 (70.7%)
Other	5 (1.8%)
Not checked	19 (6.9%)
Highest level of education, n (%)	
Elementary and/or primary school	16 (5.8%)
Secondary and/or high school	63 (22.8%)
Some college	59 (21.4%)
College degree	83 (30.1%)
Postgraduate degree	29 (10.5%)
Technical or vocational school	25 (9.1%)
Other	1 (0.4%)
BMI	
Mean (SD)	29.7 (7.2)
Median (min, max)	28.7 (16.5-57.4)
(Q1, Q3)	(24.7, 34.3)
Duration with asthma, y	
Mean (SD)	23.3 (16.8)
Median (min, max)	20.0 (1.0-66.0)
(Q1, Q3)	(9.5, 37.0)
Number of asthma exacerbations during the past 12 mo	
Mean (SD)	1.5 (2.7)
Median (min, max)	0.0 (0.0-24.0)
(Q1, Q3)	(0.0, 2.0)
≥1	135 (48.9%)
Clinician rating of asthma severity, n (%)	
Mild (1 controller)	51 (18.5%)
Moderate (2 controllers)	134 (48.6%)
Severe (3 or more controllers)	91 (33.0%)
Clinician rating of asthma control, n (%)	
Well-controlled	65 (23.6%)
Not well-controlled	105 (38.0%)
Very poorly controlled	106 (38.4%)
FEV ₁ % predicted	
Mean (SD)	79.5 (20.9)
Median (min, max)	79.0 (25.8-165.6)
(Q1, Q3)	(65.1, 90.9)

(continued)

TABLE II. (Continued)

Characteristics	Visit 1
< 60%	47 (17.0%)
60%-80%	95 (34.4%)
> 80%	132 (47.8%)
Missing	2 (0.7%)

BMI, Body mass index; FEV₁, forced expiratory volume in one second; SD, standard deviation.

A responder definition was explored by plotting cumulative distribution function (CDF) graphs based on clinician rating of CAC, PGIC, and ACQ.

RESULTS

Study sample

A total of 276 subjects enrolled and completed baseline assessments. Demographic and clinical data are presented in Table II and showed that a broad range of asthma patients were recruited.

Item performance

The means of the individual ASD items ranged from 0.4 to 0.8 with no notable ceiling effects (all less than 25%) at visit 1. However, a large floor effect (>25%) was found, ranging from 39.3% (for PM item 7) to 67.7% (AM item 5) (Table III). A *post hoc* analysis was conducted examining the weekly floor effect defined as endorsing “no problem” (ie, no symptoms, activity limitation, or nighttime awakening) responses for 7 consecutive days (days 1 to 7); the floor effect ranged from 23.1% (PM item 7) to 44.9% (PM item 10) that was much smaller than the daily floor effect. Furthermore, the percentage of subjects who reported not experiencing any problems between days 1 and 7 was also examined, equating to the 7-day average ASD score, resulting in only 7 (6.4%) such subjects. These analyses indicated that subjects experienced different issues related to their asthma on different days and, moreover, that the single day item floor effects reflected no problems with a specific issue on a particular day, rather than a floor effect on the ASD instrument.

CFA, Rasch model analysis, and DIF

CFA confirmed a unidimensional construct for the 10 ASD items, with CFI = 0.95, RMSEA = 0.17, WRMR = 1.31, and all item standardized regression coefficients > 0.80.

Rasch model analysis also supported the unidimensionality of the ASD. All items had correctly ordered response categories, and no misfitting items were identified ($P\chi^2 > .05$ for all items). The overall fit to the Rasch model was good ($\chi^2 = 24.55$; $P = .74$). No item was identified as exhibiting DIF by gender ($P > .005$ for all items).

Reliability

A Cronbach's α of 0.94 indicated good internal consistency. Furthermore, the deletion of any individual item resulted in a decrease in the overall Cronbach's α value (range: 0.927-0.934).

The 7-day average ASD score showed good to excellent test-retest reliability, with ICC values of 0.95 for the stable group defined by PGIC, 0.93 for the stable group defined by ACQ change scores, and 0.89 for the stable group defined by the FEV₁ absolute percent change.

TABLE III. Descriptive characteristics for the asthma symptom diary at visit 1: symptom items

Item	N*	Mean $\pm$ SD	Median	Range	N (%) Floor	N (%) Ceiling	Item-total Spearman correlation
AM item 1—wheezing	260	0.6 $\pm$ 0.9	0	0.0–4.0	143 (55.0%)	3 (1.2%)	0.74
AM item 2—shortness of breath	260	0.7 $\pm$ 0.8	0	0.0–4.0	135 (51.9%)	1 (0.4%)	0.73
AM item 3—cough	260	0.7 $\pm$ 0.9	1	0.0–4.0	129 (49.6%)	2 (0.8%)	0.78
AM item 4—chest tightness	260	0.6 $\pm$ 0.8	0	0.0–4.0	149 (57.3%)	1 (0.4%)	0.72
AM item 5—night symptoms	260	0.4 $\pm$ 0.6	0	0.0–3.0	176 (67.7%)	0 (0.0%)	0.72
PM item 6—wheezing	257	0.7 $\pm$ 0.9	0	0.0–3.0	130 (50.6%)	0 (0.0%)	0.80
PM item 7—shortness of breath	257	0.8 $\pm$ 0.8	1	0.0–3.0	101 (39.3%)	0 (0.0%)	0.81
PM item 8—cough	257	0.8 $\pm$ 0.9	1	0.0–4.0	121 (47.1%)	2 (0.8%)	0.72
PM item 9—chest tightness	257	0.7 $\pm$ 0.8	0	0.0–3.0	132 (51.4%)	0 (0.0%)	0.76
PM item 10—limitations	257	0.7 $\pm$ 0.9	0	0.0–4.0	149 (58.0%)	2 (0.8%)	0.75

SD, Standard deviation.

*N differed across items owing to differing numbers of missing morning or evening diary sessions.

TABLE IV. Known-groups validity: 7-d average ASD score after visit 1† by assessment of asthma control, ACQ rating, and AQLQ12+

Criteria measures	7-d average ASD score Mean $\pm$ SD (n)	Overall F value‡ (P value)	Pairwise comparison
Clinician assessment of asthma control§			
Well-controlled	0.1 $\pm$ 0.2 (n = 61)	19.30 (<.001)	1**** 2**** 3***
Not well-controlled	0.6 $\pm$ 0.6 (n = 98)		
Very poorly controlled	0.9 $\pm$ 0.6 (n = 98)		
ACQ score			
ACQ $\leq$ 0.75	0.2 $\pm$ 0.2 (n = 63)	26.84 (<.001)	1* 2**** 3****
0.75 < ACQ < 1.5	0.4 $\pm$ 0.4 (n = 56)		
ACQ $\geq$ 1.5	0.9 $\pm$ 0.6 (n = 138)		
AQLQ12+ score			
< Median	1.0 $\pm$ 0.6 (n = 121)	38.34 (<.001)	N/A
$\geq$ Median	0.3 $\pm$ 0.3 (n = 136)		

ACQ, Asthma Control Questionnaire; AQLQ, Asthma Quality of Life Questionnaire; ASD, asthma symptom diary; N/A, not available; SD, standard deviation.

†Day 2 to day 8.

‡An analysis of covariance model adjusting age, sex and visit 1 FEV₁% predicted.

§P values *<.05, **<.01, ***<.001, and ****<.0001; 1: well-controlled vs not well-controlled, 2: well-controlled vs very poorly controlled, 3: not well-controlled vs very poorly controlled.

||P values: *<.05, **<.01, ***<.001, and ****<.0001; 1: ACQ $\leq$ 0.75 vs 0.75 < ACQ < 1.5, 2: ACQ $\leq$ 0.75 vs ACQ $\geq$ 1.5, 3: 0.75 < ACQ < 1.5 vs ACQ $\geq$ 1.5.

Construct validity

The convergent validity was good with other PRO measures, with Spearman correlations ranging from 0.36 (WPAI-Asthma Absenteeism) to 0.79 (AQLQ12+ Asthma Symptoms). Correlations with AQLQ12+ Overall, AQLQ12+ Activity Limitation, and ACQ were also high ($\rho \geq 0.70$). The correlation with FEV₁% predicted was 0.30.

The ASD was able to show significant differences between groups defined by clinician assessment of asthma control, ACQ, and AQLQ12+ ($P < .001$ in each case), providing evidence of good known-groups validity. Furthermore, all pairwise comparisons were statistically significant (Table IV).

Responsiveness

The 7-day average ASD change scores between visits 1 and 3 were significantly higher for the responder group than the nonresponder group regardless of which criterion was used to define responders (clinician rating of CAC, ACQ, and PGIC; $P < .001$ in all cases). The mean 7-day average ASD change for responders ranged from -0.2 to -0.3 , compared with -0.1 to 0.0 for nonresponders (Table V). These results showed that the 7-day average ASD score has the ability to detect a symptomatic change in these asthma subjects.

Exploratory analysis of a minimal important difference and responder definition for the ASD

Using anchor-based methods, the MIDs indicated by clinician rating of “better” asthma control, “a little better” by PGIC, and as having a decrease in ACQ (indicating improvement) of ≥ 0.5 to <0.75 were -0.2 , -0.1 , and -0.1 , respectively, based on visit 1 to visit 2 data, and -0.1 , -0.1 , and -0.2 , respectively, based on visit 2 to visit 3 data. The distribution-based methods also suggested that the MID for the 7-day average ASD scores should be between -0.1 and -0.3 . The percentage of subjects in the study sample who exceeded the MID at visit 2 ranged from 13.2% (based on 1/2 SD) to 34.6% (based on PGIC, ACQ, and SEM); the percentage of subjects who exceeded the MID at visit 3 ranged from 10.3% (based on 1/2 SD) to 31.3% (based on clinician rating, PGIC, and SEM).

The CDF plots of the change scores of the 7-day ASD average score for responder and nonresponder groups all support the use of the ASD for measuring daily asthma symptoms. The separation between the responder and nonresponder groups in terms of the percentage of patients (y-axis) reaching the change scores (x-axis) across the scale was especially marked based on ACQ change scores from visit 2 to visit 3 (see Figure 1). These

TABLE V. Responsiveness: asthma symptom diary 7-d average ASD score change from visit 1* to visit 3† by clinician rating of asthma control, ACQ, and PGIC

Criteria measures	7-d average ASD change score Mean ± SD (n)		Overall F value‡ (P value)
	Responders	Nonresponders	
Clinician rating of asthma control‡	−0.3 ± 0.4 (n = 32)	−0.1 ± 0.3 (n = 192)	9.77 (<.001)
ACQ§	−0.3 ± 0.3 (n = 54)	0.0 ± 0.3 (n = 170)	11.21 (<.001)
PGIC	−0.2 ± 0.3 (n = 65)	0.0 ± 0.3 (n = 159)	9.20 (<.001)

ACQ, Asthma Control Questionnaire; ASD, asthma symptom diary; PGIC, patient global impression of change; SD, standard deviation.

*Visit 1: day 2 to day 8.

†Visit 3: 7 d before visit 3 date.

‡Responders are defined as subjects with clinician rating of asthma control as “better” and “much better” at visit 3. Nonresponders are defined as subjects with clinician ratings of asthma control as “no change,” “worse,” or “much worse” at visit 3.

§Responders are defined as subjects with an ACQ decrease score ≥0.5 from visit 1 to visit 3. Nonresponders are defined as subjects with an increase in ACQ score or an ACQ decrease score <0.5.

||Responders are defined as subjects with PGIC ratings of “extremely better,” “moderately better,” or “a little better” at visit 3. Nonresponders are defined as subjects with all other ratings (ie, “no change” to “extremely worse”).

¶An analysis of covariance model adjusting age, sex, FEV₁% predicted, and asthma symptom diary visit 1 score.

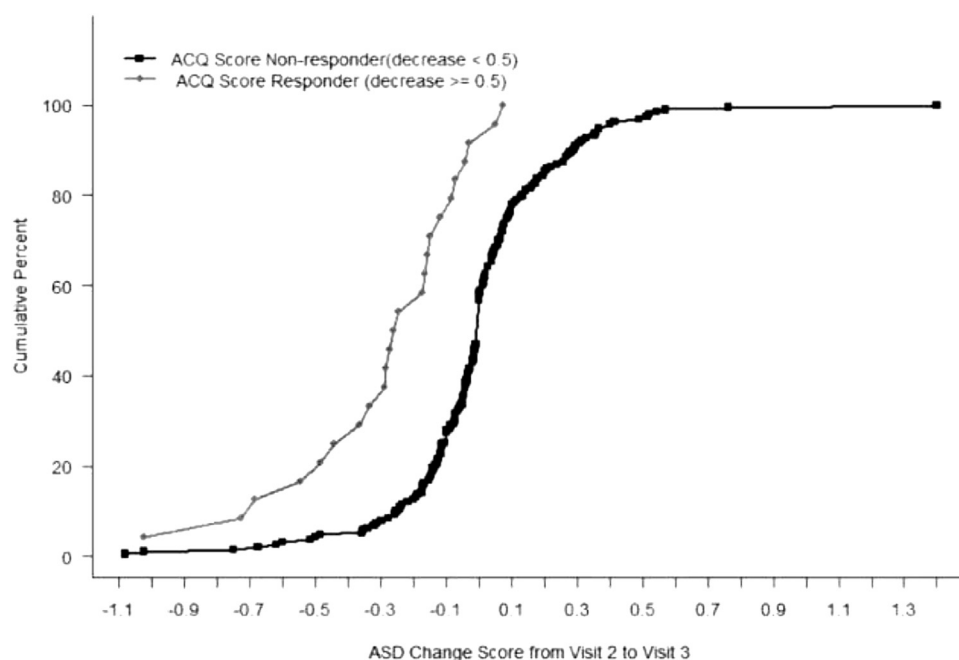


FIGURE 1. CDFs based on the ACQ score between visit 2 and visit 3. Responder defined as ACQ change scores decreased more than or equal to 0.5, vs nonresponder defined as the change score decreased less than 0.5 or the change score increased from visit 2 to visit 3. ACQ, Asthma Control Questionnaire; ASD, asthma symptom diary; CDF, cumulative distribution frequency.

CDFs supported a responder definition in the range −0.3 to −0.1 where the separation appeared to be the largest.

DISCUSSION

This was a prospective cohort study conducted to determine the psychometric properties of the ASD, as well as the performance of the individual items. The instrument was assessed for reliability, construct validity, and ability to detect a change, and a preliminary MID and responder definition were established. The findings from this study provide strong evidence of content and construct validity of the ASD, facilitating the inclusion of the measure in upcoming clinical trials.

The study population consisted of currently treated adolescent and adult asthmatic subjects across a wide range of asthma severity and sociodemographic characteristics. A range in

subjects' asthma severity (18.5% mild, 48.6% moderate, and 33.0% severe) and control (25% well-controlled, 22% not well-controlled, and 54% very poorly controlled by ACQ) was important to ensure that changes in asthma symptoms could be measured across the entire severity spectrum.

The final ASD (Figure E2, available in this article's Online Repository at www.jaci-inpractice.org) was shown to be unidimensional, a prerequisite for a summary asthma symptom score to be used in clinical trials, by both the CFA and the Rasch model analysis. DIF analysis indicated no items performed differently by gender.

The ASD is administered twice daily to capture the symptom-related impact of asthma during the night, as well as during the day. A daily score is the average of 10 items (5 morning and 5 evening items). Performance testing of the ASD daily score showed strong internal consistency and excellent reproducibility.

The morning and evening diaries also include items to evaluate the use of rescue inhaler and/or nebulizer use during the previous day or night. These items are not included in the ASD scoring, but can be used to aid the interpretation of scores.

A 7-day average ASD score is calculated when at least 4 of 7 consecutive daily scores have been captured. The 7-day average ASD score correlated highly with other relevant measures identified *a priori*, including the ACQ and AQLQ12+, and moderately with FEV₁% predicted indicating that the ASD measured the intended construct. The ASD also showed known-groups validity with significant differences in scores between groups based on clinician-rated asthma control, ACQ and AQLQ12+ scores. There is also substantial evidence that the ASD can detect a change with significant differences in 7-day average ASD scores between responders and nonresponders. MID's were calculated using both anchor-based and distribution-based models and found to be in the range of 0.1-0.3. The examination of CDF plots suggested a preliminary responder definition in the same range. Confirmatory evidence from randomized clinical trials can lend further insights into the measurement properties of the ASD. Such data may also provide the opportunity to explore further the optimum strategy for scoring the diary.

One limitation of the study is the relatively small sample size of the adolescent subjects. However, the qualitative interviews, performed in adults and adolescents, indicated that there was no difference between how adolescents and adults perceived their asthma symptoms.⁷ A future study with more adolescent participants should further support the use of the ASD in this age group. The lack of Hispanic representation in the sample, a lower percentage of males as well as the lack of subjects with a low education level also have implications for generalizability of the findings. However, the DIF analysis did not suggest that the items in the ASD performed differently in males versus females. The use of only one country in the validation phase is an additional limitation. Finally, this was a stand-alone, noninterventional observational study and, as would be typically expected, there were relatively few subjects who demonstrated symptomatic changes over the study period. Given this, the MID and responder definition analyses were exploratory and must be considered preliminary. Even though the assessment of response to treatment data is missing, the ability of the ASD to detect a change as reported by patients and clinicians has been established. Once the ASD has been implemented in clinical trials, responsiveness to a change in relation to treatment will be formally evaluated.

The ASD was devised to capture asthma symptoms comprehensively. Qualitative evidence indicates that asthma symptoms are highly variable and change frequently, and therefore a twice daily assessment, as with the ASD, is better able to capture these fluctuations. In addition, the ASD focuses on asthma symptoms and the scoring is devised to capture asthma symptoms comprehensively. The use of daily symptom diaries, preferably using an electronic mode of administration, is strongly recommended by FDA and considered desirable for asthma outcome studies sponsored by the National Institute of Health (NIH).⁶ The ASD is the first asthma diary, developed following FDA Guidance,²⁰ to undergo comprehensive psychometric evaluation and appears to be both reliable and valid in a range of patients with asthma. The ASD thus meets the needs identified by Reddel

et al¹ and the NIH asthma outcome workshop⁶ for a robust and valid measure for the daily assessment of asthma symptoms.

Acknowledgments

We thank Professor Juniper for approving the format of the ACQ and AQLQ used in this study. We also thank Diane Whalley, PhD, for her meticulous review and extremely helpful feedback during the final editing of the manuscript.

REFERENCES

1. 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.
2. Global strategy for asthma management and prevention, 2010. Available from: www.ginasthma.org. Accessed February 8, 2013.
3. Akinbami LJ, Moorman JE, Liu X. Asthma prevalence, health care use, and mortality: United States, 2005-2009. *National Health Stat Report* 2011;1-14.
4. Williams SA, Wagner S, Kannan H, Bolge SC. The association between asthma control and health care utilization, work productivity loss and health-related quality of life. *J Occup Environ Med* 2009;51:780-5.
5. Taylor DR, Bateman ED, Boulet LP, Boushey HA, Busse WW, Casale TB, et al. A new perspective on concepts of asthma severity and control. *Eur Respir J* 2008;32:545-54.
6. Busse WW, Morgan WJ, Taggart V, Togias A. Asthma outcomes workshop: overview. *J Allergy Clin Immunol* 2012;129:S1-8.
7. Globe G, Martin M, Schatz M, Wiklund I, Lin J, von Maltzahn R, et al. Symptoms and markers of symptom severity in asthma-content validity of the asthma symptom diary. *Health Qual Life Outcomes* 2015;13:21.
8. 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.
9. Juniper EF, Svensson K, Mork AC, Stahl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. *Respir Med* 2005;99:553-8.
10. Reilly MC, Tanner LA, Meltzer EO. 989 allergy impairment questionnaires: validation studies. *J Allergy Clin Immunol* 1996;97:430.
11. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407-15.
12. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. *BMJ* 1998;316:690-3.
13. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill; 1994.
14. Reeve BB, Fayers P. Applying item response theory modeling for evaluating questionnaire item and scale properties. In: Fayers P, Hays RD, editors. *Assessing Quality of Life in Clinical Trials: Methods of Practice*. 2nd ed. Oxford, NY: Oxford University Press; 2005. p. 55-73.
15. Gorsuch RL. *Factor Analysis*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates; 1983.
16. Muthén LK, Muthén B. *Mplus User's Guide*. 3rd ed. Los Angeles, CA: Muthén & Muthén; 2004.
17. RUMM. 2030: Rasch Unidimensional Measurement Models. Australia: RUMM Laboratory Pty Ltd; 2009.
18. Zumbo BD. *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-type (Ordinal) Item Scores*. Ottawa, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense; 1999.
19. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
20. Food and Drug Administration (FDA). Guidance for industry on patient-reported outcome measures: use in medical product development to support labeling claims. *Federal Registry* 2009;74:65132-3.
21. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 1994;47:81-7.
22. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;27:S178-89.
23. 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.

Clinician Assessment of Asthma Control at Baseline (CAC)

7. Ask the participant the following five questions about their asthma burden:
- a. During the past week, how many days did you experience asthma symptoms? Please think about asthma symptoms that you had after waking up to start your day and before going to bed.
_____Days/ week
 - b. During the past week, how many nights did you wake up because of your asthma symptoms?
_____Nights/week
 - c. During the past week, how many days did you use your rescue inhaler for worsening asthma symptoms?
_____Days
 - d. If during the past week you used your rescue inhaler all 7 days, how many times a day did you use it on most days?
_____Times/Day
 - e. During the past week, how much did your asthma symptoms limit your ability to perform your usual activities?
 - ☐ Not limited at all
 - ☐ Some limitation
 - ☐ Extremely limited

Using the participant's responses to questions 7a to 7d, select the corresponding cell for each column in Table A.

Table A. Assessment of Asthma Control²

LEVEL OF ASTHMA CONTROL			
Indicators	Well-Controlled	Not Well-Controlled	Very Poorly Controlled
Frequency of Asthma Symptoms	≤ 2 days/week ☐	> 2 days/week ☐	Daily ☐
Frequency of Nighttime Awakenings due Asthma	$< 1/$ week ☐	$\leq 3x/$ week ☐	$\geq 4x/$ week ☐
Interference with Activity due to Asthma Symptoms (Participant rating to item 7d)	None ☐	Some Limitation ☐	Extremely limited ☐
Frequency of use short-acting beta2-agonist use for symptom	≤ 2 days/week ☐	> 2 days/week ☐	3 or more times/ day ☐

FIGURE E1. Clinician assessment of asthma control at baseline.

Asthma Symptom Diary TM**Morning questions**

1. How severe was your wheezing from the time you went to bed last night until you woke up this morning?
 - 0 – No wheezing at all
 - 1 – Mild
 - 2 – Moderate
 - 3 – Severe
 - 4 – Very Severe

2. How severe was your shortness of breath from the time you went to bed last night until you woke up this morning?
 - 0 – No shortness of breath at all
 - 1 – Mild
 - 2 – Moderate
 - 4 – Severe
 - 5 – Very Severe

3. How severe was your cough from the time you went to bed last night until you woke up this morning?
 - 0 – No cough at all
 - 1 – Mild
 - 2 – Moderate
 - 3 – Severe
 - 4 – Very Severe

4. How severe was your chest tightness from the time you went to bed last night until you woke up this morning?
 - 0 – No chest tightness at all
 - 1 – Mild
 - 2 – Moderate
 - 3 – Severe
 - 4 – Very Severe

5. How often did you wake up because of your asthma during the night?
 - 0- Did not wake up
 - 1- Woke up 1 or 2 times
 - 2- Woke up 3 or 4 times
 - 3- Woke up 5 times or more
 - 4- Unable to sleep because of asthma

Asthma Symptom Diary TM © 2012 Amgen Inc. All rights reserved.**FIGURE E2.** Asthma Symptom Diary TM © 2012 Amgen Inc. All rights reserved.

Asthma Symptom Diary™

Evening questions

1. How severe was your wheezing from the time you woke up this morning until now?

- 0 – No wheezing at all
- 1 – Mild
- 2 – Moderate
- 3 – Severe
- 4 – Very Severe

2. How severe was your shortness of breath from the time you woke up this morning until now?

- 0 – No shortness of breath at all
- 1 – Mild
- 2 – Moderate
- 3 – Severe
- 4 – Very Severe

3. How severe was your cough from the time you woke up this morning until now?

- 0 – No cough at all
- 1 – Mild
- 2 – Moderate
- 3 – Severe
- 4 – Very Severe

4. How severe was your chest tightness from the time you woke up this morning until now?

- 0 – No chest tightness at all
- 1 – Mild
- 2 – Moderate
- 3 – Severe
- 4 – Very Severe

5. How much did your asthma symptoms limit your activities from the time you woke up this morning until now?

- 0 – Not at all
- 1 – A little
- 2 – Moderately
- 3 – Quite a bit
- 4 – Extremely

Asthma Symptom Diary™ © 2012 Amgen Inc. All rights reserved.

FIGURE E2. (CONTINUED).

TABLE E1. Details of each clinical visit

	Baseline visit	Second visit [‡]	Final visit [§]
Clinician completed			
Baseline clinical form*	X		
Change in asthma control (CAC)		X	X
Spirometry	X	X	X
Patient completed			
Written informed consent	X		
Sociodemographic Questionnaire	X		
PGIC		X	X
AQLQ12+	X	X	X
ACQ	X	X	X
WPAI-Asthma	X	X	X
ASD	X [†]	X [†]	
Diary Burden Questionnaire			X

ACQ, Asthma Control Questionnaire; AQLQ12+, Asthma Quality of Life Questionnaire; ASD, asthma symptom diary; PGIC, patients' global impression of change (in severity of asthma symptoms since the last study visit); WPAI-Asthma, Work Productivity and Activity Impairment Questionnaire-Asthma.

*Baseline clinical form included the clinician global impression of change (CGIC).

†Participant completes the daily electronic diary at home in the morning and evening.

‡The second visit was scheduled 14–18 d after completion of the baseline visit.

§The final visit was scheduled 14–18 d after completion of the second visit.

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