

PHARMACOTHERAPY

An Electronic Diary Is Shown to Be More Reliable than a Paper Diary: Results from a Randomized Crossover Study in Patients with Persistent Asthma

ANDREA M. IRELAND, PH.D., M.P.H.,¹ INGELA WIKLUND, PH.D.,^{2,*} RAY HSIEH, M.S.,¹ PETER DALE, M.SC.,³ AND ERIN O'ROURKE, B.S.¹

¹Outcomes Research, United BioSource Corporation, Bethesda, MD, USA.

²Outcomes Research, United BioSource Corporation, London, UK.

³Global Health Outcomes, GlaxoSmithKline, Uxbridge, UK.

Objectives. Test–retest reliability of an asthma paper diary versus an electronic diary (e-diary) with an integrated peak flow meter was investigated. The equivalence of the two modes was also evaluated. **Methods.** Prospective, randomized crossover study design in adolescents (12–17 years) and adults (≥ 18 years). Key inclusion criteria were persistent asthma, Asthma Control Test (ACT) scores ≥ 16 , use of inhaled corticosteroid with or without long-acting beta-agonist for ≥ 12 weeks, nocturnal awakenings < 2 times in the past week, and activity limitations < 1 per week. Participants were randomized to either paper then e-diary or e-diary then paper, to be completed for 14 days each. **Results.** Forty-seven participants completed all study visits. Weekly percentage of symptom-free days (SFDs) and rescue-free days (RFDs) were calculated. Intraclass correlation coefficients (ICCs) of Week 1 mean SFD and RFD (test) and Week 2 mean SFD and RFD (retest), respectively, were estimated in three groups defined as stable: (i) minimal changes in asthma symptoms, as measured by the global patient reported symptom change question, (ii) less than 15% change (absolute value) in 1 second FEV₁ at adjacent study visits, and (iii) changes in ACT scores less than three points for each mode. SFD demonstrated acceptable ICC (≥ 0.70) using all three definitions of asthma stability for both modes. **Conclusion.** Acceptable reproducibility of the percentage of RFD (ICC = 0.78) was only observed for the e-diary using the FEV₁ stability criterion. The ICCs for SFD and RFD were acceptable, 0.84 and 0.70, respectively, suggesting better reliability for the e-diary.

Keywords asthma, crossover design, electronic devices, patient-reported outcomes, symptoms

BACKGROUND

Asthma is a chronic condition characterized by widespread, but variable, airflow obstruction that is at least partly reversible either spontaneously or with treatment (1). It is a symptom-driven disease in which airway inflammation, in combination with bronchial hyperresponsiveness, causes recurrent episodes of dyspnea (shortness of breath), wheezing, chest tightness, and coughing (2). Because of the episodic nature of the disease, there is considerable variability in asthma symptoms (3). Asthmatic patients represent the best informants of the frequency and severity of symptoms, as well as their impacts on sleep quality, the ability to perform routine activities, and the need for rescue medications to treat worsening symptoms (4, 5).

In asthma clinical trials, documenting changes in patient-reported daytime and nighttime symptoms, and use of rescue medications is integral in order to evaluate treatment efficacy (6). Twice-daily asthma symptom diaries are most commonly used because they permit trial participants to document episodes of asthma-associated morbidity within time frames that are proximal to the actual occurrence of the events of interest (5, 7).

Historically, asthma symptom diaries have been developed for self-administration using the paper-and-pencil mode of administration (8, 9). Symptom data captured using paper diaries may be less valid due to suboptimal patient adherence with diary completion as well as the “parking lot” effect, the phenomenon where participants retrospectively complete diary entries oftentimes before study visits (5).

Electronic platforms have been increasingly used to capture patient-reported outcomes (PROs) across various indications, including asthma (10–12). By using electronic data capture (ePROs), several advantages over paper-based modes of administration are being offered, including the inclusion of alarms to remind patients to complete diaries within specified windows, minimization of missing data since devices are programmed to prevent the skipping or exclusion of diary items, as well as time stamps of patient responses. The migration of a paper-and-pencil diary to an electronic platform typically warrants minor-to-moderate changes in the content and layout of diary instructions, items, and response options. Hence, regulatory authorities also require empirical evidence, demonstrating the equivalence in psychometric performance of ePROs relative to the original, paper-and-pencil-administered PRO (13, 14).

The objectives of this study were twofold: (i) to evaluate the test–retest reliability (reproducibility) of an asthma symptoms and rescue medication use diary via the paper-based and electronic modes in clinically stable adolescent and adult participants with persistent asthma and (ii)

*Corresponding author: Ingela Wiklund, PH.D., United BioSource Corporation, 26–28 Hammersmith Grove, London, W6 7HA, UK; Tel: +44-20-8834-0100; Fax: +44-20-8834-9555; E-mail: ingela.wiklund@unitedbiosource.com

to examine the measurement equivalence of these modes of administration.

METHODS

Study Design

This was a prospective, randomized two-group crossover study. The study sample was recruited from clinics in the United States. Potential participants were identified by study sites from their database of patients willing to participate in clinical studies and contacted during a scheduled study visit or via telephone to be screened for eligibility based on the inclusion and exclusion criteria. Participants randomized to Study Arm 1 completed the twice-daily asthma symptoms and rescue medication use diary in the paper-and-pencil mode at home for 14 consecutive days, starting on the evening of the baseline visit. Participants in Arm 1 returned to the clinical site with their completed paper diaries approximately 2 weeks after baseline (Visit 2) and were crossed over to diary completion using the AM3[®] (eResearchTechnology GmbH, Germany), an integrated electronic diary (e-diary) with peak flow meter, for 14 consecutive days, starting on the evening of Visit 2. A final visit occurred approximately 2 weeks after Visit 2, where participants returned their assigned devices and site personnel uploaded diary data onto study-specific laptops. By contrast, participants in Study Arm 2 completed the diary migrated to the AM3[®] (electronic) between Visits 1 and 2, and the paper-and-pencil diary from Visit 2 to the final visit. The acceptable window of deviation for both follow-up visits was 3 days (Figure 1).

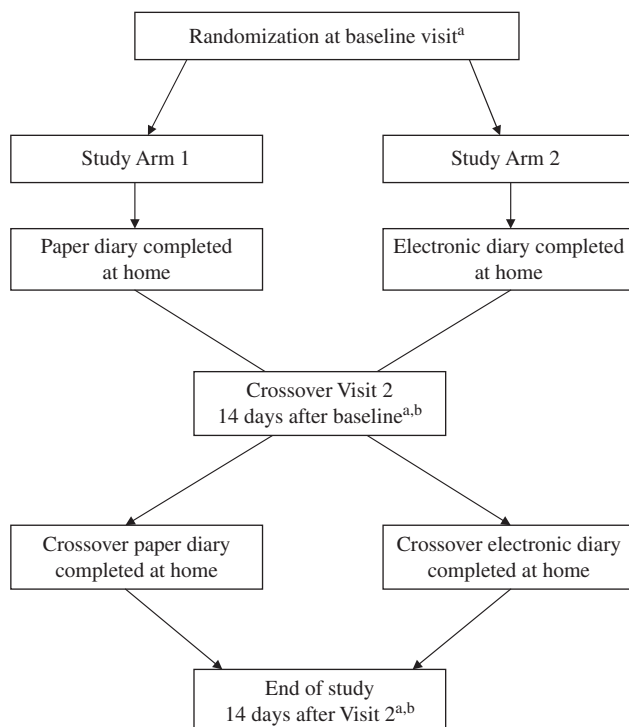


FIGURE 1.—Study design.

^aSpirometry performed in accordance with ATS/ERS guidelines; ^bwindow of deviation of 3 days was acceptable for follow-up visits.

Four asthma and allergy clinical sites in the United States (Alexandria, Virginia; Birmingham, Alabama; Blue Bell, Pennsylvania; and Wheaton, Maryland) recruited and screened participants for study eligibility. The study protocols, including recruitment and screening procedures, were reviewed and approved by the Ethics Review Committee and Institutional Review Board.

Sample

Individuals who met the following criteria were included in the study: (i) adolescents (12–17 years, inclusive) or adults (≥ 18 years); (ii) those who had clinician-diagnosed asthma, as defined by the American Thoracic Society (ATS) (15) for at least 6 months prior to study screening; (iii) those who had mild, moderate, or severe persistent asthma (16); (iv) those with total scores of ≥ 16 on the self- or telephone interview-administered versions of the Asthma Control Test (ACT) (17, 18) at study screening; (v) those who used an inhaled corticosteroid (ICS) with or without long-acting beta-agonist (LABA) at a stable dose for at least 12 weeks prior to study screening; and (vi) those who reported activity restrictions and nocturnal awakenings due to asthma not exceeding once and twice in the past week, respectively. Individuals who met any of the following criteria were deemed ineligible for the current study: (i) those who had comorbid chronic pulmonary disease or any other condition affecting lung function; (ii) those who reported use of short-acting beta-agonists (SABAs) to relieve asthma symptoms more than 3 times within the past 7 days of screening; (iii) those who completed more than one asthma-related emergency department (ED) visit in the past 3 months of screening; (iv) those who reported at least one asthma-related hospitalization in the past 6 months of screening; (v) those who had acquired bacterial or viral infections of the lower or upper respiratory tract within 4 weeks of study screening; (vi) those who had suboptimal adherence to prescribed ICS regimens based on clinician judgment; (vii) those who initiated oral corticosteroids for the treatment for worsening asthma in the past 4 weeks of screening; (viii) current smoker; and (ix) those who had cognitive impairment that would interfere with completing self-administered questionnaires.

Study Procedures

Study Visits. At baseline, eligible participants who provided written informed consent and/or assent (adolescents only) completed a brief questionnaire to capture demographic characteristics (age, gender, and race/ethnicity) as well as recent asthma symptom history and asthma resource use in the past year. Participants were then randomized to study arms to determine which mode of diary administration they would complete first. The site-stratified randomization scheme was generated for each participant number using Statistical Analysis Software (SAS[®]), version 9.3.1 (SAS Institute Inc., Cary, NC, USA). Sites were provided with sealed envelopes that contained index cards with the participants' study arm. Following successful participant enrollment, the site coordinator or designee opened a randomization envelope

containing the study ID number to assign each participant to a study arm. Deviations from the randomization scheme were not permitted.

Spirometry was performed as the final assessment at baseline and repeated at all subsequent study visits in accordance with the guidelines outlined by ATS/European Respiratory Society (ERS) (19). At Visit 2, participants returned the completed diaries from the first mode of administration and were crossed over to the second mode of diary administration. At the final study visit, participants returned the diary in the second mode and completed the self-administered version of the ACT. Indicators of potential changes in asthma status, such as the initiation of oral corticosteroids, urgent asthma care utilization, or changes in ICS and/or rescue medication after baseline, were monitored and documented by site personnel at both Visits 2 and 3.

At-home Assessments. Participants completed two additional questionnaires between study visits: the ICS Medication Adherence Questionnaire and the Patient Global Impression of Change (PGIC) in asthma status.

Measures

Asthma Symptoms and Rescue Medication Use Diary. The asthma symptoms and rescue medication use diary consisted of morning and evening questions. Participants completed the morning questions upon awakening for the day and prior to taking their morning doses of ICS. Participants were instructed to select the response that best described the impact of cardinal asthma symptoms (chest tightness, shortness of breath, cough, and wheeze) on sleep quality in the last 12 hours. The five response options to this question were “No symptoms during the night,” “Symptoms causing me to wake up once (or wake early),” “Symptoms causing me to wake twice or more (including waking early),” “Symptoms causing me to be awake for most of the night,” and “Symptoms so severe that I did not sleep at all.”

By comparison, the evening questions evaluated asthma symptom impacts on usual activities. Using a recall period of the last 12 hours, participants reported that they had either “No symptoms during the day,” “Symptoms for one short period during the day,” “Symptoms for two or more short periods during the day,” “Symptoms for most of the day, which did not affect my normal daily activities,” “Symptoms for the most of the day, which did affect my normal daily activities,” or “Symptoms so severe that I could not go to work or school or perform normal daily activities.” Participants completed the evening questions before retiring for bed at night and prior to taking evening doses of ICS. Of note, both morning and evening assessments include an item to record the total number of puffs of rescue medications taken in the previous 12 hours.

The paper and electronic (AM3[®]) versions of the asthma diary were similar, yet several differences must be noted. First, the instructions of the e-diary did not appear on the same screen as the questionnaire items or response options. Second, the response scale for the symptom items (both morning and evening questions) was

presented horizontally, whereas in the paper diary, it was vertical. The total number of puffs of rescue medication used in the previous 12 hours required a quantitative response. In the paper diary, participants wrote in those values; however, while completing the e-diary, participants selected the correct response by “togglng” to the appropriate integer. Hence, the migration of the paper diary to the AM3[®] represented a moderate level of modification, as described in the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) ePRO Task Force guidelines (13).

A pilot study was conducted to evaluate the usability of the asthma symptoms and rescue medication use diary migrated to the AM3[®] from the patient’s perspective. The findings from this unpublished study demonstrated that the aforementioned changes did not affect patient comprehension, perceived ability to respond to diary items, or ease of use. Finally, the morning and evening questions of the e-diary were only available for response from 2 am to 1:59 pm and 2:00 pm to 11:59 pm (in the participant’s time zone), respectively. The electronic (AM3[®]) version of the asthma diary included acoustic alarms to remind the patient to complete her/his diary; the alarms sounded from 8:30 am to 9:30 am for the morning session and from 9:00 pm to 10:00 pm for the evening session. At the end of each session, subjects were given an option to go back and change their answers, if required. Participants completing the paper diaries were instructed to complete the morning and evening questions within the same time frames but there was no formal mechanism to discourage deviations in the timing of diary completion.

Participant Training. Site personnel provided hands-on training experience using tutorial AM3[®] devices and patient-friendly handouts. The device manufacturer (ERT; Peterborough, Cambridgeshire, UK), together with the study team, developed the handout to train participants to operate the AM3[®] as well as to complete the diary. After the completion of this tutorial, participants independently simulated the completion of the morning and evening questions. Similarly, site personnel reviewed participants’ instructions inserted in the inside cover of the booklet containing the paper-and-pencil diary and other PROs. Site personnel, together with participants, completed supplementary diary worksheets, which outlined the dates when each diary was to be completed.

Asthma Control Test. The ACT is a validated PRO for self- and telephone interview-administration to assess the degree of asthma control (17, 18, 20). Using a recall period of the past 4 weeks, participants reported the frequency of activity limitation due to asthma, shortness of breath, nighttime awakenings, rescue medication use, and the perception of overall controlled-asthma weeks on a 5-point Likert rating scale, where “1” represented frequent symptoms or limitations due to asthma symptoms and “5” indicated no symptoms/limitations. Participant ratings on the five items were summed to yield scores ranging from 5 to 25, where higher scores were indicative of better asthma control (20). ACT scores were also used to classify

participants into one of three categories: (i) well-controlled asthma (ACT scores ≥ 20), (ii) not well-controlled asthma (scores 16–19), and (iii) very poorly controlled asthma (scores ≤ 15). Persons with changed scores at the final visit less than the absolute value of three relative to baseline, the established minimally important difference (MID) of the ACT (21), were classified as not having clinically significant changes in asthma control.

PGIC in Asthma Status

Patient perceptions of current asthma status relative to the previous week were captured using a two-item PGIC (22). The PGIC was incorporated into the study in order to quantify the extent of stability, improvement, or deterioration in asthma symptoms and rescue medication use on the final days of diary completion, independent of mode of administration. Participants could describe their current asthma symptoms as “Much better,” “Somewhat better,” “A little better,” “The same,” “A little worse,” “Somewhat worse,” or “Much worse.” Individuals who described their asthma symptoms as “A little better,” “The same,” or “A little worse” were considered as stable with respect to asthma symptoms. Similarly, response options for perceived changes in rescue medication use were “Much less often,” “Somewhat less often,” “A little less often,” “The same,” “A little more often,” “Somewhat more often,” or “Much more often.” Persons who used rescue medication with the same frequency or a little less/more often were categorized as having minimal changes in rescue medication use.

Inhaled Corticosteroid Adherence Questionnaire

A four-item ICS adherence questionnaire was adapted from the Morisky scale in order to monitor weekly changes in ICS use as participants completed the asthma symptom and rescue medication use diaries (23). The Morisky scale had been previously adapted to assess self-reported medication adherence in asthma patients (24, 25). Participants indicated how often in the past 7 days they forgot to take ICS, had been careless about taking ICS, stopped taking ICS, and used less ICS than prescribed on a 5-point Likert scale, where “1” represented “Most of the time” and “5” represented “None of the time.” Higher scores were indicative of better adherence. Total adherence scores are the mean ratings on the four items and higher scores are suggestive of better asthma maintenance medication adherence.

Statistical Analyses

Scoring of Diary. Symptom-free days (SFDs), based on the 24-hour period, were calculated from participant ratings on the morning and evening questions of the diary. If the participant reported having “no symptoms during the night” and “no symptoms during the day,” he or she was categorized as having an SFD; otherwise, it was a non-SFD. Of note, if a participant had missing data for both the morning and evening symptom questions, their data were coded as missing for that day. Similarly, if data were missing for one question, it was also coded as missing unless the participant endorsed having significant

symptoms with the other symptom question. A similar approach was undertaken to calculate RFDs (24-hour period). Weekly percentages of SFDs and RFDs were calculated for each diary mode. Week 1 and Week 2 percentages of SFDs were calculated as the number of SFDs from Days 1–8 and Days 9–15, respectively, (numerator) divided by the number of days with non-missing data during the corresponding intervals (denominator), multiplied by 100. The same algorithm was applied to calculate the percentages of RFDs for Weeks 1 and 2 for each mode.

Lung Function Parameters. Predicted values for lung function parameters were derived from the National Health and Nutrition Examination Survey (NHANES) III reference values with race corrections when appropriate (26).

Descriptive Analyses. Descriptive statistics of participant demographic and clinical characteristics, as well as asthma-related health service utilization, were performed. Counts and proportions of participants with poorly controlled, not well-controlled, and well-controlled asthma at study screening were calculated. Descriptive statistics for lung function parameters, by study visit, were also computed. The proportion of respondents who reported changes in asthma resource use (asthma-related ED visits and hospitalizations, initiation of oral steroids, and changes in ICS or rescue medications) between (i) baseline and Visit 2 and (ii) Visit 2 and the final visit were computed.

Evaluation of Order and Mode Effects. Before calculating test–retest reliability of the outcome measures of interest, repeated general linear regression models were estimated to evaluate whether there were order and mode effects (27). Mean weekly SFDs or RFDs were the dependent variables, and covariates for both models included study week, mode of administration (binary variable), study arm (binary variable), and an interaction term for week, mode, and study arm. Statistical significance of any covariates ($p < .05$) was indicative of order and/or mode effects.

Test–retest Reliability. Intraclass correlation coefficients (ICCs) of the mean percentage of SFDs for Week 1 (test) and Week 2 (retest) for the paper-and-pencil mode as well as e-diary administration were estimated in participants meeting the following stability definitions: (i) participants who reported minimal changes in asthma symptoms, as measured by the symptom item of the PGIC, (ii) those who had less than the absolute value of 15% change in forced expiratory volume in 1 second (FEV₁) at the study visit when diary mode was returned, relative to the visit where the diary was assigned, and (iii) those who did not experience clinically significant changes in asthma control (e.g., Changes in ACT scores less than the absolute value of three points). Similarly, the ICCs for the mean RFDs for Week 1 and Week 2 were calculated using the same stability definitions except the rescue medication item of the PGIC-classified participants with minimal changes in rescue medication use. ICCs equal to or greater than 0.70 are indicative of acceptable test–retest reliability (28).

Measurement Equivalence. ICCs were estimated to evaluate the equivalence of mean weekly SFD and RFD variables across the paper-and-pencil and e-diary modes of administration. For these analyses, the stable group comprised participants who had less than 15% (absolute value) change in FEV₁. Acceptable measurement equivalence was determined from ICCs between the modes of administration. ICCs greater than 0.70 were indicative of equivalence (13).

All statistical analyses were performed on SAS/STAT[®] version 9.1.3 (SAS Institute Inc.). All statistical tests used a two-sided significance level of 0.05 unless otherwise noted.

RESULTS

Disposition of Participants

Seventy-five individuals were screened for study eligibility, of which one-third ($n = 25$) were screen failures. The primary reasons for exclusion from the current study were

(i) smoking status (current smoker or recent history of smoking), (ii) use of SABA to relieve asthma symptoms more than 3 times in the past 7 days of study screening, and (iii) no use of ICS for 12 weeks or longer, prior to study screening. Fifty participants provided written informed consent and/or assent and completed the baseline visit. Forty-nine participants completed Visit 2; one Study Arm 2 participant was lost to follow-up after the baseline visit. Two participants (one from each study arm) completed Visit 2 but did not crossover to the second mode of diary administration nor complete the final visit. A total of 47 participants completed all three study visits.

Descriptive Analyses

The mean age for the sample was 36.5 ± 17.5 years (Table 1). Table 1 also summarizes characteristics of the sample in both study arms. Almost two-thirds (66%) of the sample were females and the majority of participants described their race as White (62%). Of note, other racial

TABLE 1.—Baseline sociodemographic and clinical characteristics.

Characteristic	Study Arm 1 ($n = 25$)	Study Arm 2 ($n = 25$)	Overall sample ($n = 50$)
Mean age, years (SD)	36.4 ± 20.0	36.5 ± 15.1	$36.5 (17.5)$
Female participants, n (%)	17 (68.0)	16 (64.0)	33 (66.0)
Adolescent participants, n (%)	8 (32.0)	5 (20.0)	13 (26.0)
Race, n (%)			
White	14 (56.0)	17 (68.0)	31 (62.0)
African-American	10 (40.0)	7 (28.0)	17 (34.0)
Asian	2 (8.0)	2 (8.0)	4 (8.0)
Other ^a	2 (8.0)	0	2 (4.0)
Mean age at clinician-confirmed diagnosis of asthma, years (SD)	$14.4 (13.6)$	$16.9 (14.0)$	$15.7 (13.7)$
Current controller medication ^b , n (%)			
Corticosteroids	6 (24.0)	9 (36.0)	15 (30.0)
Montelukast	6 (24.0)	6 (24.0)	12 (24.0)
Combination products—ICS + LABA	21 (84.0)	16 (64.0)	37 (74.0)
Not specified	1 (4.0)	0 (0.0)	1 (2.0)
Asthma symptom severity level ^c , n (%)			
Mild persistent	5 (20.0)	6 (24.0)	11 (22.0)
Moderate persistent	17 (68.0)	17 (68.0)	34 (68.0)
Severe persistent	3 (12.0)	2 (8.0)	5 (10.0)
Asthma control based on scores on the ACT ^d , n (%)			
Not well-controlled asthma	4 (16.0)	6 (24.0)	10 (20.0)
Well-controlled asthma	21 (84.0)	19 (76.0)	40 (80.0)
Experienced asthma symptoms, past week, n (%)			
Shortness of breath	11 (44.0)	13 (52.0)	24 (48.0)
Chest tightness	5 (20.0)	5 (20.0)	10 (20.0)
Wheezing	5 (20.0)	3 (12.0)	8 (16.0)
Coughing	5 (20.0)	8 (32.0)	13 (26.0)
Mean FEV ₁ predicted, L (SD)	$2.8 (0.9)$	$2.8 (0.6)$	$2.8 (0.8)$
Mean number of missed school/work days, past week (SD)	$0.0 (0.0)$	$0.0 (0.0)$	$0.0 (0.0)$
Mean number of asthma-related emergency department visits, past year (SD)	$0.1 (0.3)$	$0.1 (0.4)$	$0.1 (0.3)$
Mean number of asthma-related hospitalizations, past year (SD)	$0.0 (0.2)$	$0.0 (0.0)$	$0.0 (0.1)$
Comorbid conditions, n (%)			
Allergies	24 (96.0)	24 (96.0)	48 (96.0)
Eczema	5 (20.0)	5 (20.0)	10 (20.0)
Esophageal reflux or gastroesophageal reflux	5 (20.0)	10 (40.0)	15 (30.0)
Hay fever	8 (32.0)	11 (44.0)	19 (38.0)
Rhinitis	13 (52.0)	14 (56.0)	27 (54.0)
Sinusitis	5 (20.0)	8 (32.0)	13 (26.0)

Note: ^aIndividuals who self-reported their race/ethnicity as Native Hawaiian, Pacific Islander, American-Indian, or Alaskan Native were collapsed into the "Other" category due to small sample size.

^bParticipants could select more than one response. The total percentages may be greater than 100.

^cAsthma symptom severity levels were assessed by clinicians in accordance with the guidelines outlined in the National Asthma Education and Prevention Program Report.

^dPersons with ACT scores ≥ 20 were labeled as having "well-controlled asthma"; those with scores between 16 and 19, inclusive, had "not well-controlled asthma."

groups, including African-Americans (34%) and Asians (8%), were also represented within the sample (participants could self-report more than one race/ethnicity). Clinician-confirmed diagnoses of asthma occurred, on average, when participants were adolescents. Slightly more than two-thirds (68%) of the sample had moderate persistent asthma. Four out of every five participants had well-controlled asthma. Shortness of breath was the most common cardinal asthma symptom experienced by participants within the past week of baseline (48%), followed by coughing (26%). The mean baseline FEV₁ was 2.79 ± 0.78 L (range: 1.16–4.93 L).

Asthma resource use within the past year was extremely low in this sample as indicated by the reported number of asthma-related ED visits and hospitalizations. Further, participants did not experience any school or work limitations due to asthma in the week preceding the visit. Almost all participants (96%) reported comorbid allergies; however, concomitant conditions were also prevalent within this sample, including rhinitis (54%), hay fever (38%), esophageal reflux or gastroesophageal reflux (30%), sinusitis (26%), and eczema or atopic dermatitis (20%).

Participants were adherent to the ICS medications throughout the study, with median ICS medication adherence scores at Weeks 1 and 2 for each mode of administration of 5.0 (the highest overall score), thereby suggesting that sample participants were compliant with their prescribed asthma controller medications. No study participant reported urgent, asthma-related visits between the baseline visit and Visit 2. However, one participant initiated a round of oral steroids and modified their ICS dose between Visit 2 and the final visit. Mean lung function parameters at Visit 2 and the final visit were similar to those at baseline (mean FEV₁: Visit 2, 2.78 ± 0.79 L (range: 1.16–4.93); final visit, 2.75 ± 0.78 L (range: 1.29–4.95)).

Diary Compliance

Interesting trends were observed with respect to compliance to diary completion. Daily compliance rates with morning questionnaire administered as a paper-based

diary ranged from 96% to 100%, whereas this range was lower with the AM3[®] (76–96%). Similar trends were observed for the evening questionnaire (range: paper, 96–100%; AM3[®], 56–96%).

Outcome Variables

Table 2 summarizes the mean percentages of SFDs and RFDs by study arm and week within each mode of diary completion. Mean weekly SFDs ranged from 49.8 ± 35.0% (Week 1 of Study Arm 2) to 71.0 ± 26.9% (Week 2 of Study Arm 1). Similarly, mean weekly RFDs ranged from 78.7 ± 28.6% (Week 1 of Study Arm 2) to 89.5 ± 17.9% (Week 1 of Study Arm 1).

Evaluation of Order and Mode Effects

In the mixed models for repeated measures, no significant study arm, mode of diary administration, or interaction terms were observed for SFDs and RFDs. In light of these findings, diary data were pooled across study arms for test–retest reliability analyses.

Test–Retest Reliability

The ICCs for SFD for both the paper-and-pencil and electronic modes among participants who reported having minimal changes in asthma symptoms, as measured by the PGIC, exceeded 0.75 (ICC: paper, 0.77; electronic, 0.75), thereby demonstrating acceptable test–retest reliability (Table 3). By contrast, the ICCs for RFDs among persons with stable rescue medication as determined by the PGIC were somewhat lower (paper, 0.60; electronic, 0.63). Among persons who had less than the absolute value of 15% change in FEV₁ between consecutive study visits, the reproducibility of the SFDs using paper-and-pencil and electronic modes was also acceptable (ICC: paper, 0.74; electronic, 0.71). The ICCs for RFDs for the e-diary met the criterion for acceptable test–retest reliability, having an ICC of 0.78; however, this result was not observed for the paper diary (ICC = 0.67).

Among the 47 participants who completed all study visits, 11 participants experienced clinically meaningful

TABLE 2.—Mean percentages of SFD and RFD by study arm and weeks.

Study arm	Mean SFD (SD)			
	First diary mode of administration		Second diary mode of administration	
	Week 1 (n)	Week 2 (n)	Week 1 (n)	Week 2 (n)
Study Arm 1 (paper then electronic)	71.0 (26.9) (n = 25)	68.2 (33.0) (n = 25)	64.5 (33.9) (n = 24)	70.1 (31.2) (n = 24)
Study Arm 2 (electronic then paper)	49.8 (35.0) (n = 24)	55.5 (38.5) (n = 24)	60.9 (39.9) (n = 23)	58.6 (35.6) (n = 23)
Study arm	Mean RFD (SD)			
	First diary mode of administration		Second diary mode of administration	
	Week 1 (n)	Week 2 (n)	Week 1 (n)	Week 2 (n)
Study Arm 1 (paper then electronic)	89.5 (17.9) (n = 25)	82.1 (26.3) (n = 25)	82.7 (27.7) (n = 24)	80.9 (27.8) (n = 24)
Study Arm 2 (electronic then paper)	78.7 (28.6) (n = 24)	81.2 (25.1) (n = 23)	80.9 (27.7) (n = 23)	78.9 (29.5) (n = 23)

TABLE 3.—Test–retest reliability of the percentages of SFDs and RFDs by stability definition and diary mode of administration.

	Week 1 mean	Week 2 mean	Mean difference (<i>p</i>)	Pearson correlation ^a	ICC
PGIC ^b					
Paper-and-pencil					
SFD (<i>n</i> = 37)	64.5	65.6	1.1 (0.78)	0.76	0.77
RFD (<i>n</i> = 40)	87.5	82.9	−4.6 (0.17)	0.62	0.60
AM3 [®]					
SFD (<i>n</i> = 41)	57.0	64.5	7.5 (0.05)	0.76	0.75
RFD (<i>n</i> = 41)	80.9	82.9	2.0 (0.56)	0.64	0.63
<±15% Change in FEV ₁ ^{b,c}					
Paper-and-pencil					
SFD (<i>n</i> = 44)	65.6	64.5	−1.1 (0.76)	0.73	0.74
RFD (<i>n</i> = 44)	84.9	80.4	−4.5 (0.15)	0.69	0.67
AM3 [®]					
SFD (<i>n</i> = 42)	56.3	60.6	4.4 (0.31)	0.71	0.71
RFD (<i>n</i> = 41)	83.0	80.7	−2.3 (0.39)	0.79	0.78
Nonachievement of MID for the ACT at final visit ^b					
Paper-and-pencil					
SFD (<i>n</i> = 36)	71.9	66.3	−5.5 (0.16)	0.73	0.73
RFD (<i>n</i> = 36)	88.9	82.0	−6.9 (0.06)	0.61	0.57
AM3 [®]					
SFD (<i>n</i> = 36)	61.2	69.8	8.6 (0.04)	0.73	0.70
RFD (<i>n</i> = 36)	81.5	83.4	1.8 (0.64)	0.58	0.58

Note: ^a*p* < 0.0001 for all correlations except calculated for RFD using the electronic mode of administration in participants who did not achieve the MID on the ACT at the final visit (*p* < 0.001).

^bThree definitions of stability were employed for these analyses: (1) having minimal changes in asthma symptoms or rescue medication based on the PGIC, (2) <±15% change in FEV₁, and (3) Changes in ACT scores at the final visit (relative to baseline) <±3 points.

^cThe participant who initiated the round of oral steroid and changed ICS medications during the study met the stability definition based on lung function. However, the authors decided to exclude this person from these analyses.

TABLE 4.—Score equivalence of the SFD and RFD for Weeks 1 and 2 between paper-and-pencil and AM3[®] in stable participants, as determined by <±15% change in FEV₁ between consecutive study visits.

Variable	Paper-and-pencil mean	AM3 [®] mean	Mean difference (<i>p</i>)	Pearson correlation ^a	ICC
SFD (<i>n</i> = 41)	63.3	57.9	−5.4 (0.06)	0.84	0.84
RFD (<i>n</i> = 41)	82.9	81.9	−1.1 (0.72)	0.70	0.70

Note: ^a*p* < 0.0001 for all correlations.

changes in asthma control, as indicated by the achievement of the ACT MID at the final visit. The ICCs for SFD using this stability definition for the paper-and-pencil and electronic modes were 0.73 and 0.70, respectively. The ICCs for RFDs did not meet the acceptable thresholds for either mode (paper, 0.57; electronic, 0.58).

Measurement Equivalence

The ICC between modes of administration for SFDs (Weeks 1 and 2) was high (ICC = 0.84) (Table 4), thereby suggesting equivalence between the paper-and-pencil and electronic modes of administration. The ICC for RFDs was acceptable (ICC = 0.70). Further, the *t*-tests showed no significant differences in mean scores between modes.

DISCUSSION

In a recent systematic review of randomized, double-blind, placebo-controlled trials published from 2004 to 2006, Dinan and colleagues found that symptom diaries were the most common PRO used in asthma clinical trials (29). This finding may be attributable, in part, to evidence

requirements outlined by regulatory agencies. The European Medicines Agency, for example, recommends the supplementation of lung function end points with patient-reported symptom data in order to evaluate the efficacy of asthma medications (30). As described previously, there is considerable variability in asthma symptoms. Twice-daily diaries represent an optimal approach to balance the competing demands of minimizing patient burden while simultaneously collecting reliable and valid retrospective symptom data in clinical trials (6, 14). Technological features of asthma symptom diaries migrated to electronic devices may further enhance data quality. For example, questionnaires can be programmed so that participants can complete diaries at specific time intervals. Further, alarms remind participants to complete diaries, thereby improving diary compliance.

In this study, the test–retest reliability of the percentage of SFDs, as captured by both the paper-and-pencil diary and AM3[®], met or exceeded the threshold of acceptable reproducibility with all three definitions of asthma stability. Similarly, adequate reproducibility of the percentage of RFDs was observed in the AM3[®] diary using the stability definition of the change in FEV₁ between adjacent

study visits. These findings were achieved despite the fact that the number of patients included in the analytic samples was comparatively small.

RFDs via the paper-and-pencil diaries did not yield acceptable test–retest reliability with any stability definition ($ICC < 0.67$), suggesting that the timely completion and higher data quality yielded by using the AM3[®] device delivers a more reliable assessment of RFD recordings, albeit in only one of the three groups tested, that is, stability defined as $< \pm 15\%$ change in FEV_1 . The asthma symptoms and rescue medication use diary migrated to the AM3[®] demonstrated equivalence with the mean percentage of SFDs ($ICC = 0.84$) and RFDs ($ICC = 0.70$) captured by the paper-and-pencil diary. This finding was consistent with a meta-analytic review evaluating the equivalence of paper and electronic PROs (10).

In this study, compliance to diary completion was somewhat higher for participants completing the paper-and-pencil mode relative to the diary in the AM3[®] device. While participants could complete the paper-and-pencil diaries outside of the protocol-delineated time frames (including immediately prior to study visits), participants who completed the diary on the AM3[®] device could not go back and add their entries at a later time point, as they were limited to completing the diary within the 12-hour windows in the morning and evening. While time stamps were not available for the paper diaries, to confirm our assumptions about paper diary completion patterns, in a sample of adults with chronic pain, Stone and colleagues found that while participants reported completing diaries on a regular basis as instructed, the rates of true diary adherence, as determined by photosensors, were much lower (7).

Because the purpose of the study was to assess test–retest reliability in addition to documenting the concordance between paper and electronic modes of administration, the patients included had to be in a stable phase of their condition as defined by study inclusion and exclusion criteria. Hence, even though two-thirds of the patients had moderate to persistent asthma, the study population is not representative of an average asthma population. Another potential limitation of the study was the absence of procedures to mitigate learning effects between the two modes of administration, as there was no gap between the two assessment periods and a new baseline was not reestablished. However, the evaluation of sequence and mode of administration had showed no significant study arm, mode of diary administration, or interaction terms for SFDs and RFDs. Additional psychometric evaluation must also be undertaken to examine the construct validity (cross-sectional and longitudinal) and responsiveness of the diary migrated to the AM3[®] platform.

By contrast, several study strengths should be highlighted. This study conformed with many of the recommendations outlined in the ISPOR ePRO Good Research Practices Task Force Report, including usability testing to confirm that participants were able to complete the e-diary, implementation of a randomized crossover design to evaluate measurement equivalence of the paper-and-pencil and AM3[®] diaries, and statistical testing of potential order and

mode effects (13). This study was also noteworthy because it employed three definitions of stable asthma patients for the test–retest analyses, two of which were informed by PRO measures, namely, the PGIC and ACT, even though the results provided stronger support for SFD than RFD. The precedence of using subjective participant ratings to measure changes in health states and the inclusion of a recall period that was easily understood by patients represented some of the advantages of using the PGIC to classify stable asthma patients in this study (31). Further, Schatz and colleagues documented that change scores greater than the absolute value of three were indicative of clinically significant changes in asthma status, thereby differentiating stable from unstable asthma patients (21).

In addition, slightly more than three-quarters of the sample had moderate or severe persistent asthma, thus mimicking the distribution of symptom severity levels in asthma trial populations.

Purposeful deliberation was given to the study design. For example, clinicians together with health outcome researchers collaborated to create stringent participant inclusion and exclusion criteria in order to enroll participants with active asthma who would maintain the same level of symptom frequency throughout the study. ICS adherence and urgent asthma-related resource utilization were captured to document whether participants experienced moderate-to-severe asthma exacerbations while enrolled in the study. As such, the authors were able to discern that one participant initiated a round of oral steroids and changed his or her ICS medication sometime between baseline and Visit 2.

In conclusion, the asthma symptoms and rescue medication use diary migrated to the AM3[®] have adequate test–retest reliability and measurement equivalence with the paper-and-pencil version of the diary.

DECLARATION OF INTEREST

At the time of the study, Andrea M. Ireland, Ingela Wiklund, Ray Hsieh, and Erin O'Rourke were employed by United Biosource Corporation (UBC), which provides consulting and other research services to pharmaceutical, device, government, and nongovernment organizations. Peter Dale is a former former salaried employee of GlaxoSmithKline.

In their salaried positions, Drs. Ireland and Wiklund, as well as Mr. Hsieh and Mrs. O'Rourke, work with a variety of companies and organizations. They received no payment or honoraria directly from these organizations for the services rendered.

Funding for this study was provided by Glaxo-SmithKline.

REFERENCES

1. Global Initiative for Asthma (GINA). Global strategy for asthma management and prevention. 2011. Available at: <http://www.ginasthma.org/guidelines-gina-report-global-strategy-for-asthma.html>. Accessed May 17, 2012.

2. Lemanske Jr RF, Busse WW. Asthma: clinical expression and molecular mechanisms. *J Allergy Clin Immunol* 2010; 125:S95–S102.
3. Taylor DR, Bateman ED, Boulet LP, Boushey HA, Busse WW, Casale TB, Chanez P, Enright PL, Gibson PG, de Jongste JC, Kerstjens HA, Lazarus SC, Levy ML, O'Byrne PM, Partridge MR, Pavord ID, Sears MR, Sterk PJ, Stoloff SW, Szeffler SJ, Sullivan SD, Thomas MD, Wenzel SE, Reddel HK. A new perspective on concepts of asthma severity and control. *Eur Respir J* 2008; 32:545–554.
4. Leidy NK, Revicki DA, Geneste B. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value Health* 1999; 2:113–127.
5. Stull DE, Leidy NK, Parasuraman B, Chassany O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions. *Curr Med Res Opin* 2009; 25:929–942.
6. Reddel HK, Taylor DR, Bateman ED, Boulet LP, Boushey HA, Busse WW, Casale TB, Chanez P, Enright PL, Gibson PG, de Jongste JC, Kerstjens HA, Lazarus SC, Levy ML, O'Byrne PM, Partridge MR, Pavord ID, Sears MR, Sterk PJ, Stoloff SW, Sullivan SD, Szeffler SJ, Thomas MD, Wenzel SE. 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.
7. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. *BMJ* 2002; 324:1193–1194.
8. Santanello NC, Barber BL, Reiss TF, Friedman BS, Juniper EF, Zhang J. Measurement characteristics of two asthma symptom diary scales for use in clinical trials. *Eur Respir J* 1997; 10:646–651.
9. Wyrwich KW, Ireland AM, Navaratnam P, Nolte H, Gates Jr DF. Validation of the AQLQ12+ among adolescents and adults with persistent asthma. *Qual Life Res* 2011; 20:903–912.
10. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health* 2008; 11:322–333.
11. Jiang H, Han J, Zhu Z, Xu W, Zheng J, Zhu Y. Patient compliance with assessing and monitoring of asthma. *J Asthma* 2009; 46:1027–1031.
12. Meltzer EO, Kelley N, Hovell MF. Randomized, cross-over evaluation of mobile phone vs paper diary in subjects with mild to moderate persistent asthma. *Open Respir Med J* 2008; 2:72–79.
13. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, Lenderking WR, Cella D, Basch E. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009; 12:419–429.
14. Food and Drug Administration (FDA). Guidance for industry on patient-reported outcome measures: use in medical product development to support labeling claims. *Federal Register* 2009; 74: 65132–65133.
15. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. *Am Rev Respir Dis* 1987; 136:225–244.
16. National Heart, Lung, and Blood Institute (NHLBI). Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma. National Asthma Education and Prevention Program. Bethesda, MD: National Heart, Lung, and Blood Institute (NHLBI), 2007.
17. Kosinski M, Kite A, Yang M, Rosenzweig JC, Williams A. Comparability of the Asthma Control Test telephone interview administration format with self-administered mail-out mail-back format. *Curr Med Res Opin* 2009; 25:717–727.
18. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, Murray JJ, Pendergraft TB. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol* 2004; 113:59–65.
19. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J. Standardisation of spirometry. *Eur Respir J* 2005; 26:319–338.
20. Schatz M, Sorkness CA, Li JT, Marcus P, Murray JJ, Nathan RA, Kosinski M, Pendergraft TB, Jingran P. Asthma Control Test: reliability, validity, and responsiveness in patients not previously followed by asthma specialists. *J Allergy Clin Immunol* 2006; 117:549–556.
21. Schatz M, Kosinski M, Yarlus AS, Hanlon J, Watson ME, Jingran P. The minimally important difference of the Asthma Control Test. *J Allergy Clin Immunol* 2009; 124(719–723):e711.
22. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989; 10:407–415.
23. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care* 1986; 24:67–74.
24. Erickson SR, Coombs JH, Kirking DM, Azimi AR. Compliance from self-reported versus pharmacy claims data with metered-dose inhalers. *Ann Pharmacother* 2001; 35:997–1003.
25. Ivanova JI, Birnbaum HG, Hsieh M, Yu AP, Seal B, van der Molen T, Emani S, Rosiello RA, Colice GL. Adherence to inhaled corticosteroid use and local adverse events in persistent asthma. *Am J Manag Care* 2008; 14:801–809.
26. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; 159:179–187.
27. Diggle PJ, Liang KY, Zeger SL. *Analysis of Longitudinal Data*. 1st ed. Oxford: Clarendon Press, 1994.
28. Hays R, Revicki DA. Reliability and validity (including responsiveness). In: Fayers P, Hays R, eds. *Assessing Quality of Life in Clinical Trials*. New York: Oxford University Press, 2005: 25–39.
29. Dinan MA, Compton KL, Dhillon JK, Hammill BG, Dewitt EM, Weinfurt KP, Schulman KA. Use of patient-reported outcomes in randomized, double-blind, placebo-controlled clinical trials. *Med Care* 2011; 49:415–419.
30. Committee for Medicinal Products for Human Use. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. London: European Medicines Agency, 2005.
31. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 2009; 17:163–170.

Copyright of Journal of Asthma is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.