

CORRESPONDENCE BETWEEN PAPER AND ELECTRONIC VISUAL ANALOG SCALES AMONG ADULT ASTHMATICS

Michael R. Hufford, Ph.D.¹ & Saul Shiffman, Ph.D.^{1,2}

¹ invivodata, inc.

² University of Pittsburgh

ABSTRACT

Background: The Visual Analog Scale (VAS) is a widely used assessment format that allows patients to rate a variety of subjective symptoms on a continuous scale. The typical implementation of a VAS question is a 10cm line anchored on both ends with verbal descriptors representing different extremes of patient experience (e.g., 'Not at all' and 'Extremely'). The implementation of handheld technology to the collection of patient diary data (Stone & Shiffman, 1994; Shiffman, Hufford & Paty, 2001) has introduced a new variant of the traditional paper-based VAS. Because current handhelds have screens less than 10cm in width or height, VASs on these devices are shorter than 10cm, which could affect their psychometric properties. This validation study sought to examine the correspondence between paper and electronic versions of the same 3-item VASs. **Design:** This was a randomized, crossover design with all subjects completing both paper and electronic versions of the same questionnaires in counterbalanced order. **Methods:** Forty asthma patients were recruited from newspaper and radio advertisements to complete an in-clinic assessment on a single occasion. After completing demographic and asthma severity forms, as well as multiple measures of their Peak Expiratory Flow, patients completed both paper and electronic versions of three VAS scales assessing asthma severity, and the impact of their asthma on their activity level and social activities. The VAS anchors were "Not at all" and "Extremely" for all three measures. Order of administration was counterbalanced. Based on psychometric analyses, the three VAS scores were averaged to form a reliable VAS composite (reliability = 0.84-0.90). One subject was a statistical outlier on each of the 3 VAS assessments and removed from the analyses. **Results:** The overall correlation between the paper and electronic versions of the VAS composites was 0.86. We also examined the correlations between the paper and electronic VAS scales and external criterion variables (e.g., asthma severity and quality of life assessments). Correlations with the electronic VASs were equal to or higher than those observed with paper VAS scales. **Conclusions:** This validation study confirmed that subjects rate their experience similarly regardless of whether the assessments are represented on paper or on a handheld electronic diary. More broadly, these results are in agreement with other research showing a high degree of correspondence between electronic and paper versions of the same assessments (Hank & Schwenkmezger, 1996; Jamison, et al., 2001).

INTRODUCTION

Researchers have been attempting to collect real-time data from patients in their natural environment for more than 50 years (Verbrugge, 1980). Advances in both the technology and methodology of real-time data collection have enabled researchers to capture reliable and valid momentary data from patients in the real world (Hufford, Shiffman, Paty, & Stone, 2001). Despite the many advantages associated with the use of electronic diaries (Hufford et al., 2001; Shiffman & Hufford, 2001; Stone & Shiffman, 1994; Stone et al., under review), some methodological issues are raised by asking patients to use an electronic device to enter data that had previously been entered on paper diaries. One issue is the psychometric equivalence of paper versus electronically administered patient self-report assessments. These issues arise when considering the uses of truncated electronic VASs in domains where 100mm paper VASs have previously been used.

This study sought to examine the psychometric equivalence between a 10cm paper- and truncated-electronic 3-item VAS measure for the assessment of asthma severity in a sample of adult asthmatics.

METHODS

Subjects. Forty asthmatic subjects participated in the study. Seventy-five respondents replied to local newspaper or radio advertisements regarding an asthma study. Sixty-one of these respondents completed the phone screening. To qualify for the study, subjects had to be at least 18 years of age, currently experiencing asthma, and be willing to complete the peak expiratory flow (PEF) and self-report measures as part of this study. Ten respondents did not qualify for the study. The remaining 51 respondents were stratified based on their self-reported asthma severity into one of 5 severity categories (1 = Very Mild, 2 = Mild, 3 = Moderate, 4 = Severe, 5 = Very Severe). Ten of the qualified respondents were not interested in participation. In an attempt to ensure adequate variability to allow for correlations between measures to emerge, no more than 35% of subjects were allowed to enroll in the study from any one of the asthma severity categories. One potential participant was excluded because a stratum had already been filled. One subject was a statistical outlier on each of the 3 VAS assessments and was removed from the analyses. The resulting 39 subjects completed the experimental session and were included in the analyses reported below.

Procedure. After passing the phone screening and arriving for their experimental session, subjects signed an informed consent. Next, subjects completed a demographic and asthma severity form and provided three readings of their PEF. Based on random assignment to an order condition for completion of the questionnaires, patients then completed both the paper and electronic versions of 3-item VAS measure, a validated daily quality of life (QoL) measure (Hyland & Crocker, 1995), and the SF-8-Abbreviated version (QualityMetric, Inc.). The experimental session lasted approximately 30 minutes for each subject. This report focuses on the VAS data. Data from Hyland's QoL measure and the SF-8 are used as criterion variables to evaluate the validity of the VAS ratings. Lastly, subjects completed a debriefing questionnaire regarding the electronic and paper versions of the measures, and were paid \$15 for their participation. This study was approved by a local IRB.

VAS Measures. Patients completed both paper and electronic versions of three VAS scales assessing the impact of their asthma on their activity level, social activities, and asthma severity. The VAS anchors were "Not at all" and "Extremely" for all three measures. After psychometric analyses indicating this was appropriate, the three VAS scores were averaged to form a VAS composite. The paper VAS scales are reproduced below:

Questions about the PAST 24 HOURS....

1. My asthma has affected my activity level?

Not at all

Extremely

2. My asthma has affected my social activities?

Not at all

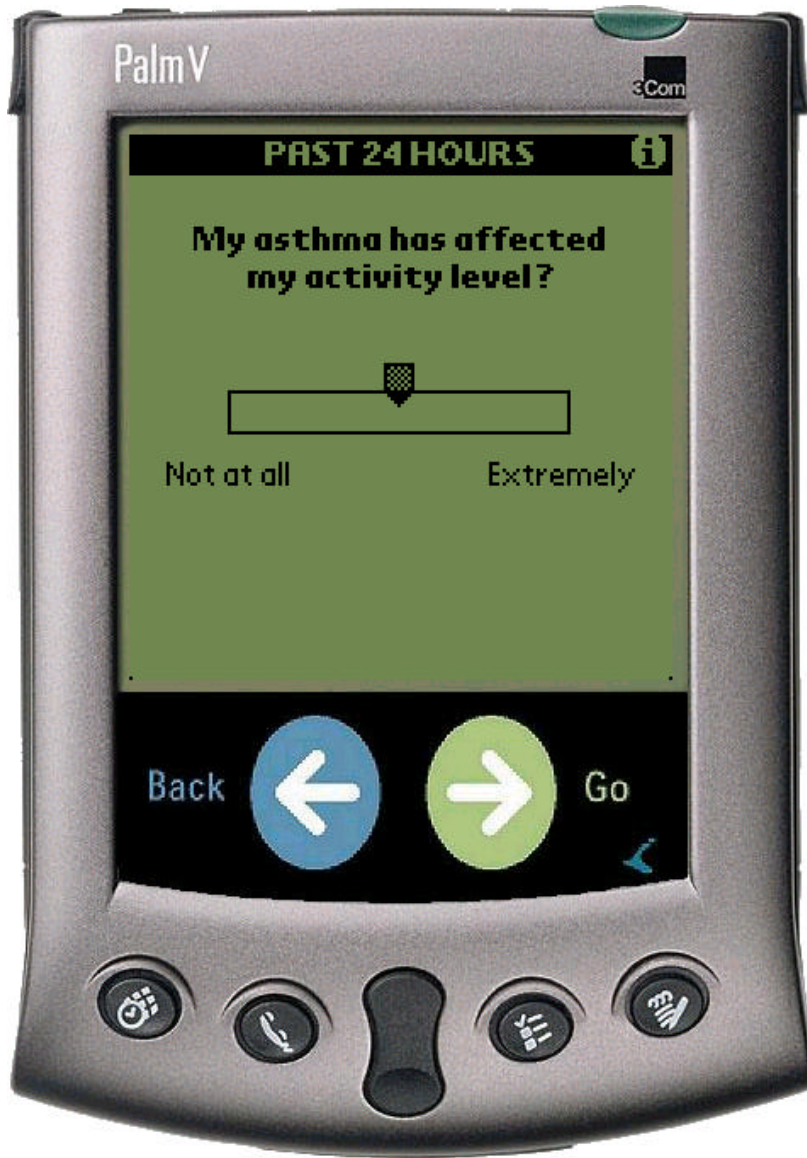
Extremely

3. Asthma Severity?

Not at all

Extremely

A sample electronic version of the VAS on the invivosystem Patient Experience Diary (PED) is reproduced below (not to scale).



RESULTS

The correlation between the electronic PED and paper VAS composite scores was 0.86.

We compared the two input methods in other ways. To assess whether the PED ratings might contain more noise (e.g., due to a smaller VAS line), we looked at several parameters. The variability in the two types of ratings was comparable. We also examined the average inter-item correlation among the three VAS ratings, an indicator of reliability. The average was slightly higher for the PED VAS (0.74 vs. 0.64), suggesting that the PED VAS was more reliable. Correspondingly, the Cronbach's alpha for the PED VAS composite was 0.90, vs. 0.84 for paper. Thus, the PED ratings appear to be slightly *less* noisy and more reliable than the paper VAS ratings. The mean of the PED composite was reliably slightly higher than the paper VAS mean (see table below).

To examine validity, we examined correlations with external criterion assessments of constructs expected to be tapped by the VAS assessments. At screening for the study, participants had rated their asthma severity. We correlated this rating with the VAS asthma severity rating. For PED, this correlation was 0.53; for paper, it was 0.48. We also correlated the VAS asthma ratings and the VAS composite with Hyland's daily QoL measure (Hyland & Crocker, 1995). For the composite measure, the correlations were comparable (PED, = 0.60; paper = 0.61). For the asthma ratings, the correlation with PED VAS was slightly higher (PED = 0.53; paper = 0.48). Thus, VAS ratings obtained on PED showed reliability and validity at least equal to than paper VAS ratings.

Finally, participants were asked their preference for rating methods. Over three quarters (77.5%) preferred making their VAS ratings on electronic PED.

| | PED | | PAPER | | Difference | | Corr |
|--|--------------|--------------|--------------|--------------|-------------|------------------|-------------|
| | Mean | SD | Mean | SD | | | |
| Asthma and activity level | 35.54 | 23.63 | 30.67 | 24.81 | 4.87 | ns | 0.69 |
| Asthma and social activities | 29.41 | 25.95 | 20.49 | 20.83 | 8.92 | p<.005 | 0.72 |
| Asthma severity | 39.21 | 24.25 | 35.31 | 23.38 | 3.90 | p<.10 | 0.85 |
| AVERAGE | 34.72 | 22.40 | 28.82 | 20.06 | 5.90 | p<.005 | 0.86 |
| Inter-item correlation | 0.74 | | 0.64 | | | | |
| Reliability (alpha) | 0.90 | | 0.84 | | | | |
| Correlation of Asthma VAS w screening ratings | | | | | | | |
| Asthma severity | 0.53 | | 0.48 | | | | |
| Correlation of VAS w Hyland asthma daily QoL ratings | | | | | | | |
| Asthma severity | 0.60 | | 0.49 | | | | |
| Composite | 0.60 | | 0.61 | | | | |

CONCLUSIONS

This validation study confirmed that subjects rate their experience similarly regardless of whether the assessments are represented on paper or on a handheld electronic diary. Because the reliability of many single-item VAS scales are unknown, it is important to note the random error of measurement sets a limit of the maximum correlation between measures¹. That is, if two measures have low reliabilities to begin with, then the correlation between the tests will be considerably lower than the value that would be obtained were the measures more reliable (Edwards, 1984).

The results showed that VAS ratings obtained on an electronic diary were at least as reliable as paper-based 100 mm VAS ratings. Electronic VAS ratings also correlated with external criterion variables at least as well as those based on paper VAS ratings. In this study, the electronic VAS produced slightly higher absolute scores on some rating scales. However, for studies analyzing VAS ratings that do not cross methods, this phenomenon, even if true, should not affect within-study comparisons.

More broadly, these results are in agreement with other research showing a high degree of correspondence between electronic and paper versions of the same assessments (e.g., Hank & Schwenkmezger, 1996; Jamison, et al., 2001).

REFERENCES

- Edwards, A. L. (1984). *An Introduction to Linear Regression and Correlation*. Second Edition. W.H. Freeman & Company, New York.
- Hank, P. & Schwenkmezger, P. (1996). Computer-assisted versus paper-and-pencil based self-monitoring: An analysis of experiential and psychometric equivalence. In J. Fahrenberg & M. Myrtek (Eds.) *Ambulatory Assessment*, pp. 85-99.
- Hufford, M. R., Shiffman, S., Paty, J., & Stone, A. A. (2001). Ecological momentary assessment: Real world, real-time measurement of patient experience. In J. Fahrenberg & M. Myrtek (Eds.) *Progress in Ambulatory Assessment*, pp. 69-92. Seattle, WA USA: Hogrefe & Huber Publishers.
- Hyland, M. E. & Crocker, G. R. (1995). Validation of an asthma quality of life diary in a clinical trial. *Thorax*, 50, 724-730.
- Jamison et al. (2001). Electronic vs. paper VAS ratings: A randomized, crossover, comparative study utilizing healthy volunteers. *The Journal of Pain (Suppl.)*, 2, 7.
- Shiffman, S. & Hufford, M. R. (2001). Subject experience diaries in clinical research, Part 2: Ecological Momentary Assessment. *Applied Clinical Trials*, 10(3), 42-48.
- Shiffman, S., Hufford, M. R., & Paty, J. (2001). Subject experience diaries in clinical research, Part 1: The patient experience movement. *Applied Clinical Trials*, 10(2), 46-56.
- Stone, A. A. & Shiffman, S. (1994). Ecological Momentary Assessment: Measuring real world processes in behavioral medicine. *Annals of Behavioral Medicine*, 16, 199-202.
- Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (under review). Collection of patient data with paper diaries: Evidence of Poor Compliance. *British Medical Journal*.
- Verbrugge, L. M. (1980). Health diaries. *Medical Care*, 18, 73-95.

¹ The correlation between X_1 and X_2 relies in part on the error term associated with each measure (ϵ_1 and ϵ_2 , respectively). Thus, r_{12} will take on its maximum value only if $\epsilon_1 = \epsilon_2 = 0$, and thus can only occur if $\epsilon_1 = \epsilon_2 = 0$.