Self-reported Adherence Measures
What Do They Assess and How Should We Use Them?

John F. Steiner, MD, MPH

Scientific papers on adherence almost universally begin by noting that nonadherence with treatment is pervasive, that it has a profound impact on health outcomes and costs, and that interventions to improve adherence are only modestly effective. These inherent difficulties in measuring and improving adherence are compounded by substantial confusion about the meaning of the term “adherence” itself. Commonly, adherence is viewed as an underlying characteristic of a person. As a result, researchers vainly seek a single “gold standard” measure of adherence.1 A more useful conceptual model proposes that adherence is a chain or cascade of linked behaviors that begins with seeking care, followed by receiving a diagnosis, instituting lifestyle changes, starting medications, filling prescriptions, and taking pills.2 In this model, measures of adherence assess discrete behaviors rather than a single underlying construct, and even small lapses in adherence at each level can lead to cumulatively large effects on clinical outcomes.3

Most adherence measures count the frequency of behaviors—they assess how often people keep appointments, fill prescriptions, or take their pills. Although self-reported measures may also quantify adherence, their accuracy is limited by recall or social desirability bias.4 Uniquely, self-reported adherence measures can provide insight into the reasons for nonadherence. Because these measures are inexpensive to administer and provide rich, clinically useful information about behavior, continued development of self-reported adherence measures remains an important task.

Self-reported adherence measures differ in the behaviors they assess, their approach to quantifying those behaviors, the time frame of assessment, and their menu of potential reasons for nonadherence. Some measures assess multiple steps in the behavioral cascade, using single items or brief multi-item scales.5,6 Others focus on a specific behavior such as medication-taking.7–9 In the 1970s, Haynes and Sackett first established the validity of a single question about medication-taking for patients with hypertension, “People often have difficulty taking their pills for one reason or another. Have you ever missed any of your pills?.”7 Pooled analysis of 4 studies using this question or minor variants showed that its sensitivity for detecting nonadherence as defined by pill count was 55% and specificity was 87%.10 Because the positive predictive value of a report of nonadherence was high (82%), these authors concluded that clinicians could confidently believe individuals who acknowledged nonadherence.10

In the early 1980s, Morisky and colleagues8,9 developed a 4-item adherence measure that has since been expanded to 8 items. The original Morisky measure assessed 4 reasons why people might fail to adhere to antihypertensive medications—forgetting to take pills, carelessness, and stopping pills because they are either feeling better or worse. In their initial validation study, the likelihood of achieving blood pressure control declined as the number of reasons for omitting medications increased.8 Another commonly used measure of adherence in hypertension, the Hill-Bone scale, includes 8 items about
medication-taking, 3 about dietary sodium use, 2 about appointment-keeping, and 1 about filling prescriptions. A recent study has raised concerns that, in German translation, both measures had substantial “ceiling effects” (ie, few individuals reported nonadherence), as well as problems with internal consistency, only fair agreement with each other, and little correlation with blood pressure control.

In this issue of Medical Care, Voils et al describe and validate a new self-reported measure of the extent of medication-taking for patients with hypertension, along with a detailed list of reasons why patients may not take their pills. The 3 “extent” items assessed whether individuals took all their doses, missed or skipped doses, or were unable to take doses of their medications over the previous 7 days. The 3-item scale had acceptable internal consistency and substantial variability, and was modestly but significantly correlated with systolic and diastolic blood pressure ($r=0.27$ for both). The scale also demonstrated convergent validity with the 8-item Morisky scale ($r=0.62$). The 21 reasons for nonadherence were largely independent of each other. Each was endorsed by 7%–27% of respondents, suggesting that these reasons would be relatively common in practice. Although these findings require confirmation in other patient populations and disease states, this is certainly a promising debut for a new measure.

Voils et al recognize that a self-reported adherence measure must navigate between the simplicity that clinicians value and the reliability and predictive validity required by researchers. Their measure has several appealing features. The 3 questions about the extent of nonadherence are likely to identify nonadherent individuals more reliably than a single item, yet remain brief enough to incorporate into a clinical conversation. The long list of reasons for nonadherence may only need to be administered to individuals who “test positive” on the initial 3-item screen. Further, each of these reasons can trigger a targeted conversation about strategies to overcome that specific adherence barrier.

This measure also holds promise for researchers. Because it is shorter than either version of the Morisky instrument and more focused than the Hill-Bone scale, the Voils measure may be useful in telephone-based or internet-based surveys. Detailed exploration of reasons for nonadherence could be omitted, or limited to the subset of individuals who report nonadherence. For both clinical and research use, this measure should be easily programmable into electronic health records, e-health applications such as patient portals, or interactive telephone voice response tools.

As adherence measures are so often viewed as alternative tools for assessing a single underlying construct, researchers have seldom explored the simultaneous use of multiple measures to enhance detection of nonadherence. In a recent study in Medical Care, nonadherence with refills and with self-reported medication-taking (measured with the 4-item Morisky scale) were independently associated with a lower likelihood of blood pressure control. This finding suggests that concurrent use of prescription refill measures and self-reported medication-taking might increase the identification of nonadherence.

The establishment of the new Patient-Centered Outcomes Research Institute reminds us that we have both the imperative and the capacity to include the patient perspective in practice and research. Researchers can help accomplish this goal by developing tools that incorporate structured, patient-reported adherence information into electronic health records, internet-based patient portals, and other communication tools. If further research supports the early promise of the adherence measure proposed by Voils et al, their instrument will be a strong candidate for inclusion in such efforts.

REFERENCES

Initial Validation of a Self-Report Measure of the Extent of and Reasons for Medication Nonadherence

Corrine I. Voils, PhD,*† Matthew L. Maciejewski, PhD,*† Rick H. Hoyle, PhD,‡ Bryce B. Reeve, PhD,§ Patrick Gallagher, PhD,* Christopher L. Bryson, MD,*∥¶ and William S. Yancy Jr, MD, MHS*†

Background Self-report measures of medication nonadherence confound the extent of and reasons for medication nonadherence. Each construct is assessed with a different type of psychometric model, which dictates how to establish reliability and validity.

Objectives: To evaluate the psychometric properties of a self-report measure of medication nonadherence that assesses separately the extent of nonadherence and reasons for nonadherence.

Research Design: Cross-sectional survey involving the new measure and comparison measures to establish convergent, discriminant, and predictive validity. The new measure was readministered 2–21 days later.

Subjects: A total of 202 veterans with treated hypertension were recruited from the Durham Veterans Affairs Medical Center.

Measures: A new self-report measure assessed the extent of nonadherence and reasons for nonadherence. Comparison measures included self-reported medication self-efficacy, beliefs about medications, impression management, conscientiousness, habit strength, and an existing nonadherence measure.

Results: Three items assessing the extent of nonadherence produced reliable scores for this sample, α = 0.84 (95% confidence interval, 0.80–0.87). Correlations with comparison measures provided evidence of convergent and discriminant validity. Correlations with systolic (r = 0.27, P < 0.0001) and diastolic (r = 0.27, P < 0.0001) blood pressure provided evidence of predictive validity.

Reasons for nonadherence were assessed with 21 independent items. Intraclass correlations were 0.58 for the extent score and ranged from 0.07 to 0.64 for the reasons.

Conclusions: The dual conceptualization of medication nonadherence allowed a stronger evaluation of the reliability and validity than was previously possible with measures that confounded these 2 constructs. Measurement of self-reported nonadherence consistent with psychometric principles will enable reliable, valid evaluation of interventions to reduce nonadherence.

Key Words: adherence, reliability, scale development, self-report

Medication nonadherence is a significant clinical problem in chronic disease management. Medication nonadherence is associated with increased health care spending, hospitalization rates, morbidity, and premature mortality.

Obtaining accurate estimates of medication nonadherence is essential to determine where intervention resources should be directed. There is no “gold standard” for assessing nonadherence. Pill refills, pill counts, and computerized bottle caps can approximate how much medication patients are consuming, but only patients can report reasons for not taking their medications. The self-report method is also appealing because it can be administered in any setting, is low cost, and can provide immediate feedback at the point of care.

Considerable effort has been made to assess self-reported medication nonadherence, yielding several self-report instruments. Although an expert committee recently identified medication nonadherence as one of the constructs that should be assessed routinely in electronic health records, it did not recommend an existing measure and suggested that further work was needed. Others have also concluded that the existing measures lacked reliability and validity.

Recently, we examined existing self-report measures of medication nonadherence to determine how measurement could be improved. A key limitation of existing measures is that they confound 2 related but distinct nonadherence constructs: the extent to which doses are missed and the reasons for missing doses. Each construct is assessed by a different type of psychometric model, which has important measurement implications (Table 1). Existing self-report measures were not developed with this distinction in mind,
TABLE 1. Using Effect Versus Causal Indicators for Measuring Extent of Nonadherence and Reasons for Nonadherence, Respectively

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Extent of Nonadherence (Effect Indicators)</th>
<th>Reasons for Nonadherence (Causal Indicators)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. items</td>
<td>Few</td>
<td>Many</td>
</tr>
<tr>
<td>Interitem correlations</td>
<td>High, positive</td>
<td>No requirement</td>
</tr>
<tr>
<td>Reliability</td>
<td>Internal consistency (Cronbach coefficient-x), test-retest (same reference period, stable population)</td>
<td>Test-retest</td>
</tr>
<tr>
<td>Validity</td>
<td>Content (focus groups, cognitive interviews); construct (factor analysis); convergent and discriminant validity (correlations with other measures)</td>
<td>Content (focus groups, cognitive interviews)</td>
</tr>
<tr>
<td>Scoring</td>
<td>Averaged to estimate level of nonadherence</td>
<td>Items stand alone as descriptors for reasons for nonadherence</td>
</tr>
<tr>
<td>Use in research</td>
<td>Covariate or stratifying variable for treatment efficacy; outcome in intervention to improve adherence</td>
<td>Provide qualitative information about reasons for nonadherence; help inform the design or tailoring of interventions</td>
</tr>
<tr>
<td>Generic vs. disease specific</td>
<td>Generic—to be used in any disease population and for any medication</td>
<td>Both—some reasons are reported across populations, whereas others are more relevant for a particular disease population or medication</td>
</tr>
</tbody>
</table>

Participants

Participants were recruited from the Durham Veterans Affairs Medical Center, where Institutional Review Board approval was obtained. Veterans older than 40 years with a diagnosis of HTN were identified from electronic medical records. Inclusion criteria determined during a screening telephone call were: prescription of at least 1 antihypertensive medication, stability of antihypertensive regimen for at least 3 months prior, and receipt of antihypertensive medications from Durham Veterans Affairs Medical Center. Exclusion criteria were: cognitive impairment based on a 6-item screener, unable to complete questionnaires unaided, unable to communicate in English or by telephone, resident in a nursing home or receiving home health care, or health problem that precludes participation.

Recruitment and Study Procedures

Patients meeting initial inclusion criteria received a recruitment letter and telephone call. Eligible patients were scheduled for an assessment, which coincided with a scheduled medical appointment when possible. Reminder letters were mailed before the visit and included instructions not to smoke or consume caffeine or alcohol for at least 60 minutes before the assessment.

At time 1, the research assistant conducted the consent process, obtained 2 BP readings according to recommended standards,18 and administered the self-report measures orally. At time 2, the research assistant obtained 2 BP readings and administered the new nonadherence measure orally. Participants received $20 for each assessment.

Measures

At time 1, we collected demographic data, medication names, and dosing instructions in addition to the measures listed below.

Extent of Nonadherence

Participants rated the extent to which they have missed doses of their medications over the past 7 days by 5 items with response options: strongly disagree, strongly agree, neutral, agree, and strongly agree. Higher scores indicate greater levels of nonadherence.
Reasons for Nonadherence
Participants rated 23 reasons for missed antihypertensive medications in the past 7 days on 5-point scales anchored by not at all and very much (see Data, Supplemental Digital Content 1, http://links.lww.com/MLR/A332 which includes the extent and reasons items). Higher scores indicate greater endorsement of each reason for missing dose(s).

Self-Efficacy to Take Antihypertensive Medication
Self-efficacy to take medication as prescribed was assessed with the 13-item Medication Adherence Self-Efficacy Scale-Revised,15 in which participants rated how sure they are that they can take their medication in certain situations. The internal consistency (estimated using Cronbach coefficient-α, hereafter referred to as α) in our sample was 0.94.

Beliefs About Medicines
The 18-item Beliefs about Medicines Questionnaire (BMQ)19 comprises of 4 factors: beliefs about the necessity of prescribed medication (Specific-Necessity; α = 0.84); beliefs about the danger of dependence and long-term toxicity and the disruptive effects of medication (Specific-Concerns; α = 0.78); beliefs that medicines are harmful, addictive poisons that should not be taken continuously (General-Harm; α = 0.74); and beliefs that medicines are overseen by doctors (General-Overuse; α = 0.80). Higher scores indicate more positive beliefs about medications for the Necessity factor, and lower scores indicate more positive beliefs about medications on the others.

Impression Management
Participants completed the 20-item Impression Management subscale of the Balanced Inventory of Desirable Responding20 to assess the tendency to portray a positive perception to others (α = 0.72).

Conscientiousness
Participants indicated how well 25 adjectives described them from very inaccurate to very accurate. This scale was designed to assess facets of conscientiousness. Items were presented in a different random order for each participant to minimize order effects. Fifteen of the items yielded 3 facets: orderliness (α = 0.84), impulse control (α = 0.63), and reliability (α = 0.80). The remaining 10 items not contributing to a facet in an exploratory factor analysis were not analyzed.

Habit Strength
The habit strength measure was adapted from previous measures21,22 to assess the frequency of medication-taking behavior with 1 item and situational consistency with 5 additional items (physical location, time of day, people present, mood, and event). Habit strength was calculated as the product of frequency and the mean of the 5 situational consistency items.

BP
All BP measurements were performed using a digital sphygmomanometer. Participants sat quietly for 5 minutes. Arm circumference was measured for the appropriate size cuff, and BP was measured while participants were sitting in a chair with back rested, both feet on the floor, and arm supported at heart level. A second reading was obtained 1 minute later. The average of the 2 BP readings was used in analyses.

Existing Measure of Self-Reported Nonadherence
Participants completed the 8-item version of the Morisky scale.7 Two items (2 and 5) assess the extent of nonadherence over 2 weeks and yesterday, and 6 items assess reasons for nonadherence: forgetting (items 1, 4, and 8), feeling worse (item 3), BP in control (item 6), and feeling hassled (item 7). A summary score was calculated.7 This scale was administered to provide evidence of convergent validity of the extent measure and to gain initial evidence as to whether separating extent from reasons improves the relationship with the criterion (BP).

Analysis
The extent of nonadherence is represented by an effect indicator model, in which a person’s level of medication nonadherence determines their item (indicator) responses.23 Descriptive statistics were examined for all extent items. Excessive skewness is indicated by values >2 and excessive kurtosis by values >7.24 Internal consistency and confirmatory factor analysis (CFA) were conducted to examine the extent to which a single latent variable contributed to indicators of extent of nonadherence. CFA is an inferential test of the hypothesis that a single factor accounts for the data. Because of the small item set, a 2-factor model was not tested. To set the metric for the CFA model, all factor loadings were freely estimated, and the latent factor’s variance was set to 1. Because of the non-normal item distributions, the response scales were treated as ordinal by specifying categorical variable type. Incremental model fit was assessed by the Comparative Fit Index (CFI)25 and Tucker-Lewis Index (TLI)26; values >0.95 are generally accepted as good model fit.27 Absolute model fit was assessed by Weighted Root Mean Square Residual (WRMR), for which values <0.90 are generally accepted as good fit.28 Extent items were considered for elimination due to excessive missingness, skewness, or kurtosis; insufficient interitem and item-total correlations or factor loadings (<0.40)29; improvement in α once an item was deleted; or redundancy with other items, as indicated by interitem correlations and factor loadings much higher than those for other item pairs. The mean of all retained items was calculated and used in all analyses (see Data, Supplemental Digital Content 2, http://links.lww.com/MLR/A333 which includes the item response theory analyses).

Pearson correlations were computed between extent and comparison measures to provide information about convergent and discriminant validity. As an a priori guideline, any correlation >0.50 was considered as evidence of convergent validity and any correlation <0.30 as evidence of discriminant validity. Because self-efficacy is a proximal
determinant of behavior, we expected a large correlation between extent and medication self-efficacy. We also expected a large correlation between extent and the Morisky scale to provide evidence of convergent validity because some Morisky items assess missed doses. We expected small correlations between extent and the 4 BMQ subscales, impression management, conscientiousness, and habit strength to provide evidence of discriminant validity. Because factors other than nonadherence contribute to elevated BP, we expected a small to moderate correlation between BP and extent to provide evidence of predictive validity.

Reasons for nonadherence are represented by a causal indicator model, in which each reason for nonadherence stands alone as a descriptive indicator for the construct because they would not necessarily be correlated. Participants indicating any nonadherence (score ≥ 2 on any extent item) were considered nonadherent, and descriptive statistics for their reasons items were examined (see Figure, Supplemental Digital Content 3, http://links.lww.com/MLR/A335 which includes the histograms). Reasons items were expected to be skewed and/or kurtotic and not highly intercorrelated. We determined a priori that reasons items highly correlated (r>0.60) with one another would be examined for possible redundancy. Alpha was not calculated because it is inappropriate for causal indicators.

To examine the test-retest reliability of the nonadherence measures, intraclass correlations (ICC) were calculated for the extent summary score and each reason using a 2-way mixed model with time as a fixed variable and participants as a random variable. Test-retest reliability assumes the nonadherence constructs are stable over the 2 assessment periods (which ranged from 2 to 21 d). For participants whose reference periods for the assessment points do not align, there is concern about the stability of the nonadherence constructs, especially for the reasons scale as reasons for missing a medication may change over time. Thus, we are neither confident that test-retest reliability alone is an accurate indicator of reliability, nor should it be used alone as the only indicator to retain or remove an item from the scale.

The target sample size was 200, which provides >99.5% power to detect the significance of a correlation of ≥ 0.30 at P=0.05. The CFA was performed using Mplus (version 6). The remaining analyses were performed using SAS (version 9; SAS Institute, Cary, NC).

RESULTS

A total of 740 recruitment letters were mailed, for which 566 patients were contacted by telephone. Of those, 100 were ineligible, 210 refused, 6 died before contact, and 250 were scheduled for a time 1 visit. Of the 250, 48 did not show for their appointment, leaving a time 1 n of 202 (36%). Of those, 186 (92%) returned for the time 2 visit.

Participants were 64 years on average, of mixed racial composition, and primarily male (Table 2). Nearly three quarters of participants reported some education beyond high school, and only 10% reported insufficient income to pay bills.

### Extent of Nonadherence

Although the means of all extent items except item 4 were below the scale midpoint, the distributions were not highly skewed or kurtotic (Table 3). All items except item 4 had adequate item-total and interitem correlations. A single factor accounted for the shared variance among the 5 extent items, χ² = 25.61, P = 0.0001, CFI = 0.99, TLI = 0.99, WRMR = 0.60. Although all items had sufficient factor loadings, the item 4 loading was lower than for the other items. Items 1 and 2 were substantially correlated (r = 0.84), suggesting redundancy between them. Item 2 was retained in favor of item 1 because its wording is more specific, and item 4 was eliminated because it measures a related but different construct (ie, being late for a dose). Items 2, 3, and 5 were averaged (unweighted) to create a final summary score, M = 1.78 (SD = 0.96). These items produced reliable scores, α = 0.84 (95% confidence interval, 0.80–0.87).

As expected, extent was highly correlated with medication self-efficacy (Table 4). Although correlations between extent and the harm subscale of the BMQ and habit strength were >0.30, they were not so high as to indicate measurement of a similar construct. Correlations between extent and the necessity, concerns, and overuse subscales of the BMQ, impression management, and the 3 facets of conscientiousness were small in magnitude, demonstrating discriminant validity. Finally, predictive validity was evidenced by correlations between extent and BP: for systolic, r²02 = 0.27, P < 0.0001 and for diastolic, r²02 = 0.27, P < 0.0001.

### Reasons for Nonadherence

Means of the reasons items were well below the scale midpoint, and several distributions were positively skewed and kurtotic (Table 5). Sixty percent (N = 122) of participants were considered nonadherent (≥ 2 on any extent item). The
reason endorsed (score ≥ 2) most was I forgot (27%). The least commonly endorsed reasons were feeling too ill to take them (7%) and going on a long car/plane/bus ride (7%).

As expected, interitem correlations were, with few exceptions, sufficiently small to suggest that these items are capturing independent reasons for nonadherence (range, /C0 0.01 to 0.81, average \( r = 0.28 \)). Item 17 (I felt well) was eliminated because it was highly correlated with and subsumed by item 13 (I felt I did not need them). Item 23 (I was going on a long car/bus/plane ride) was eliminated because it was highly correlated with and subsumed by item 14 (I was traveling). The result was a list of 21 relatively independent reasons for nonadherence (range, /C0 0.01 to 0.64, average \( r = 0.28 \)).

**TABLE 4. Correlations Between Extent of Nonadherence and Individual Difference Measures (n = 202)**

<table>
<thead>
<tr>
<th>Individual Difference Measure</th>
<th>Extent Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication self-efficacy</td>
<td>r = −0.42, P &lt; 0.0001</td>
</tr>
<tr>
<td>Beliefs about medications—necessity</td>
<td>r = −0.13, P = 0.07</td>
</tr>
<tr>
<td>Beliefs about medications—concerns</td>
<td>r = 0.25, P &lt; 0.0001</td>
</tr>
<tr>
<td>Beliefs about medications—harm</td>
<td>r = 0.31, P &lt; 0.0001</td>
</tr>
<tr>
<td>Beliefs about medications—overuse</td>
<td>r = 0.16, P = 0.02</td>
</tr>
<tr>
<td>Impression management</td>
<td>r = −0.11, P = 0.12</td>
</tr>
<tr>
<td>Conscientiousness: orderliness</td>
<td>r = −0.11, P = 0.12</td>
</tr>
<tr>
<td>Conscientiousness: impulse control</td>
<td>r = −0.11, P = 0.11</td>
</tr>
<tr>
<td>Conscientiousness: reliability</td>
<td>r = −0.15, P = 0.04</td>
</tr>
<tr>
<td>Habit strength</td>
<td>r = −0.39, P &lt; 0.0001</td>
</tr>
</tbody>
</table>

**DISCUSSION**

Our self-report measure reflects a dual conceptualization of nonadherence. Although related, the 2 facets are distinct, requiring different measurement approaches and
TABLE 5. Characteristics of Reasons for Nonadherence Items

<table>
<thead>
<tr>
<th>Reason for Nonadherence</th>
<th>M (SD)</th>
<th>Skewness*</th>
<th>Kurtosis*</th>
<th>N (%) Endorsed</th>
<th>ICC (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I was busy</td>
<td>1.37 (0.93)</td>
<td>2.7</td>
<td>7.2</td>
<td>22 (18.0)</td>
<td>0.27 (0.09 to 0.43)</td>
</tr>
<tr>
<td>2. There was no one to remind me</td>
<td>1.59 (1.18)</td>
<td>2.0</td>
<td>2.8</td>
<td>33 (27.0)</td>
<td>0.47 (0.32 to 0.60)</td>
</tr>
<tr>
<td>3. They caused some side effects</td>
<td>1.25 (0.84)</td>
<td>3.3</td>
<td>10.0</td>
<td>12 (9.8)</td>
<td>0.37 (0.20 to 0.52)</td>
</tr>
<tr>
<td>4. I worried about taking them for the rest of my life</td>
<td>1.52 (1.23)</td>
<td>2.3</td>
<td>3.6</td>
<td>23 (18.9)</td>
<td>0.57 (0.43 to 0.68)</td>
</tr>
<tr>
<td>5. They cost a lot of money</td>
<td>1.34 (0.98)</td>
<td>2.9</td>
<td>7.2</td>
<td>16 (13.1)</td>
<td>0.46 (0.31 to 0.60)</td>
</tr>
<tr>
<td>6. I came home late</td>
<td>1.30 (0.75)</td>
<td>2.7</td>
<td>6.5</td>
<td>22 (18.0)</td>
<td>0.34 (0.16 to 0.49)</td>
</tr>
<tr>
<td>7. I did not have any symptoms of high blood pressure</td>
<td>1.27 (0.77)</td>
<td>3.3</td>
<td>11.2</td>
<td>18 (14.8)</td>
<td>0.24 (0.06 to 0.41)</td>
</tr>
<tr>
<td>8. I was with friends or family members</td>
<td>1.22 (0.67)</td>
<td>3.6</td>
<td>14.0</td>
<td>16 (13.1)</td>
<td>0.23 (0.04 to 0.39)</td>
</tr>
<tr>
<td>9. I was in a public place</td>
<td>1.21 (0.67)</td>
<td>3.4</td>
<td>10.9</td>
<td>14 (11.5)</td>
<td>0.32 (0.14 to 0.48)</td>
</tr>
<tr>
<td>10. I was afraid of becoming dependent on them</td>
<td>1.37 (1.08)</td>
<td>2.8</td>
<td>6.1</td>
<td>14 (11.5)</td>
<td>0.50 (0.35 to 0.63)</td>
</tr>
<tr>
<td>11. I was afraid they may affect my sexual performance</td>
<td>1.54 (1.28)</td>
<td>2.1</td>
<td>2.9</td>
<td>21 (17.2)</td>
<td>0.58 (0.45 to 0.69)</td>
</tr>
<tr>
<td>12. The time to take them was between my meals</td>
<td>1.28 (0.84)</td>
<td>3.2</td>
<td>10.6</td>
<td>16 (13.1)</td>
<td>0.29 (0.12 to 0.45)</td>
</tr>
<tr>
<td>13. I felt I did not need them</td>
<td>1.22 (0.73)</td>
<td>4.0</td>
<td>16.6</td>
<td>14 (11.5)</td>
<td>0.64 (0.51 to 0.74)</td>
</tr>
<tr>
<td>14. I was traveling</td>
<td>1.20 (0.70)</td>
<td>4.0</td>
<td>16.8</td>
<td>13 (10.7)</td>
<td>0.31 (0.14 to 0.47)</td>
</tr>
<tr>
<td>15. I was supposed to take them more than once a day</td>
<td>1.20 (0.75)</td>
<td>4.2</td>
<td>17.2</td>
<td>10 (8.2)</td>
<td>0.07 (0.11 to 0.25)</td>
</tr>
<tr>
<td>16. I had other medications to take</td>
<td>1.30 (0.91)</td>
<td>3.3</td>
<td>10.2</td>
<td>15 (12.3)</td>
<td>0.63 (0.51 to 0.73)</td>
</tr>
<tr>
<td>17. I felt well</td>
<td>1.36 (1.00)</td>
<td>2.9</td>
<td>7.3</td>
<td>18 (14.8)</td>
<td>0.46 (0.31 to 0.60)</td>
</tr>
<tr>
<td>18. They make me want to urinate while away from home</td>
<td>1.48 (1.17)</td>
<td>2.3</td>
<td>3.9</td>
<td>21 (17.2)</td>
<td>0.44 (0.28 to 0.58)</td>
</tr>
<tr>
<td>19. I ran out of medication</td>
<td>1.34 (0.98)</td>
<td>3.0</td>
<td>8.1</td>
<td>16 (13.1)</td>
<td>0.14 (0.03 to 0.31)</td>
</tr>
<tr>
<td>20. I was afraid the medication would interact with other medication I take</td>
<td>1.34 (1.01)</td>
<td>3.1</td>
<td>8.1</td>
<td>15 (12.3)</td>
<td>0.54 (0.39 to 0.66)</td>
</tr>
<tr>
<td>21. My blood pressure was too low</td>
<td>1.17 (0.64)</td>
<td>4.1</td>
<td>17.2</td>
<td>10 (8.2)</td>
<td>0.15 (0.03 to 0.33)</td>
</tr>
<tr>
<td>22. I was feeling too ill to take them</td>
<td>1.16 (0.67)</td>
<td>4.6</td>
<td>21.4</td>
<td>9 (7.4)</td>
<td>0.51 (0.36 to 0.64)</td>
</tr>
<tr>
<td>23. I was going on a long car/bus/plane ride</td>
<td>1.16 (0.67)</td>
<td>4.6</td>
<td>21.4</td>
<td>9 (7.4)</td>
<td>0.58 (0.45 to 0.69)</td>
</tr>
</tbody>
</table>

*Among 122 participants reporting any degree of nonadherence on at least 1 extent of nonadherence item.

*Defined as responding at least 2 on 1–5 scale where 1 = not at all and 5 = very much.

*n = 113 due to data missing for time 2.

CI indicates confidence interval, ICC, intraclass correlation.

Evaluation using different psychometric models. Consistent with our conceptualization, the extent of nonadherence was assessed with 3 positively correlated items, which produced reliable scores and correlated as expected with related constructs. Also consistent with our conceptualization, reasons for nonadherence were assessed with several independent items. In contrast to the extent items, which are averaged into an overall score, the reasons items are treated individually in a descriptive manner to inform treatment decisions or to tailor interventions to increase adherence.

Although short measures are always desirable for research and clinical settings, reliability and validity must be a priority. Multiple items (in this case, 3), rather than a single item, are needed to assess the extent of nonadherence because the reliability, and therefore predictive validity, of effective indicators is increased by the use of multiple items. Some portion of the unreliability in individual items is not shared with other items that vary in wording and, as such, the composite score, which reflects the commonality across items, reduces the impact of unreliability on scores.

The need to assess the reasons for nonadherence with many items to provide construct validity is underscored by the endorsement of individual items by only a small proportion of participants and by the lack of sizable interitem correlations. Future research is needed to determine whether the improved content validity translates to better ability to detect intervention outcomes or improves clinical practice.

One practical advantage of measuring extent and reasons separately is that the measurement process can be streamlined: The 3 extent items can be used to help identify patients with suboptimal levels of adherence, followed by the reasons items to identify targets of intervention if necessary. This approach is similar to depression screening, where a positive 2-item Patient Health Questionnaire is followed by a more comprehensive depression measure or diagnostic interview.33 Although the response scale for extent items is continuous, various cutoffs could classify nonadherent individuals, depending on a researcher’s or clinician’s goals.

Another advantage of measuring the 2 constructs separately is that longitudinal assessments may provide a more fine-grained picture of medication-taking behavior. Previous studies have suggested that adherence is episodic.4 Because of the content of existing self-report measures, it is difficult to determine whether the extent of nonadherence or the reasons for nonadherence vary over time. The separate measurement of extent and reasons in this self-report measure enables independent assessment of both the constructs. The ICCs for extent individual items and total score were moderate to large over 2–21 days, suggesting that some individuals report a consistent level of nonadherence. In contrast, ICCs for many reasons items were more modest, indicating that some reasons for nonadherence are highly variable. More research is needed to elucidate longitudinal patterns of medication-taking behavior, which could inform intervention development and clinical practice.

As scale development is an iterative and ongoing process, we will continue to refine the measure and build the body of evidence for its reliability and validity. For example, because the extent items and response scales were not subjected to cognitive interviews, cognitive interviews should be conducted to improve further the instructions and evaluate different response scales. More evidence of convergent
validity of the extent measure could be obtained by comparing it to electronic medication monitoring, commonly characterized as a more objective method. More evidence of criterion-related validity is needed as well. Although the extent and Morisky measures were highly correlated, only the extent measure was significantly associated with BP. Thus, despite some shared variance between the extent and the Morisky scales, they account for different variance in BP. This pattern of correlations should be examined in other samples to provide more evidence on the predictive validity of the new measure. Criterion-related validity may be difficult to establish in HTN because BP is highly variable and reflects the influence of other factors. In diseases in which the outcome is more stable over time, such as low-density lipoprotein cholesterol treated by statins, a higher correlation might be expected between extent and the criterion. Equally important is establishing the psychometric properties of this measure in other diseases and patient populations. As expected of effect indicator models, the extent measure should have stable psychometric properties across diseases and populations. The reasons measure will need to be tailored to characteristics of the disease, patient populations, and medications used.

In summary, the dual conceptualization of the extent of nonadherence and reasons for nonadherence provided a framework for the measurement of these constructs. Using this conceptualization, we developed a preliminary version of a measure to assess these facets separately, thereby allowing a stronger evaluation of the reliability and validity than was possible with existing measures. By improving the measurement of self-reported nonadherence, we hope to enable better evaluation of interventions to improve patient-centered outcomes and clinical practice.

REFERENCES