Validation of a New Three-Item Self-Report Measure for Medication Adherence

Ira B. Wilson[1], **Yoojin Lee**[1], **Joanne Michaud**[1], **Floyd J. Fowler Jr.**[2], and **William H. Rogers**[3]

Ira B. Wilson: ira_wilson@brown.edu

[1]Department of Health Services, Policy & Practice, Brown University School of Public Health, Box G-121-7, 121 South Main St., Providence, RI 02912, USA

[2]Center for Survey Research, University of Massachusetts, Boston, MA, USA

[3]Center for Health Solutions, Tufts Medical Center, Boston, MA, USA

Abstract

Few self-report measures of medication adherence have been rigorously developed and validated against electronic drug monitoring (EDM). Assess the validity of the 3-item self-report scale by comparing it with a contemporaneous EDM measure. We conducted an observational study in which adherence assessments were done monthly for up to 4 months for 81 patients with HIV who were taking antiretroviral medications. We report results for both HIV antiretroviral medications, and also for other, non-HIV-related medications. Raw and calibrated self-report adherence measures, electronic drug monitoring adherence measures, and sociodemographic variables. The mean age of patients was 46 years, 37 % were female, 49 % had some education beyond high school, 22 % were Black, and 22 % were Hispanic. Cronbach's alphas for the 3-item scale for HIV and non-HIV medications were 0.83 and 0.87, respectively. The mean differences (raw/uncalibrated self-report scale minus EDM) for HIV and non-HIV medications were 7.5 and 5.2 points on a 100-point scale (p < 0.05 for both). Pearson correlation coefficients between the calibrated 3-item scale and the EDM for HIV and non-HIV medications were 0.47 and 0.59, respectively. The c-statistics for the ROC curves for the calibrated scale, using cut-offs of 0.8 and 0.9 for the EDM gold standard measure to define non-adherence, were between 0.74 and 0.76 for HIV and non-HIV medications. This 3-item adherence self-report scale showed good psychometric characteristics and good construct validity when compared with an EDM standard, for both HIV and non-HIV medications. In clinical care it can be a useful first-stage screener for non-adherence. In clinical research and quality improvement settings it can be a useful tool when more complex and expensive methods such as EDM or pharmacy claims are impractical or unavailable.

Keywords

Medication adherence; Highly active antiretroviral therapy; Self-report; HIV; Patient compliance

Introduction

Medication adherence researchers have long debated the validity of self-report [1]. A wide variety of self-report measures have been used, but few have been carefully tested. Care providers could use a valid self-report measure to screen patients for potential adherence problems. Researchers could also benefit from valid self-report measures. While there are a variety of other ways to assess medication adherence, all are complex and therefore somewhat expensive [2]. Alternative methods include the use of pharmacy claims databases [3], use of electronic drug monitoring devices [4], and the use of unannounced counts [5]. Unannounced pill counts can be done both in person, and over the telephone.

We previously reported the development and field testing of three self-report items [6, 7]. To develop these items, we conducted a literature search to identify self-report items that had been used for either HIV antiretroviral medications or for other medications. An example is the commonly used Morisky scale [8]. After four rounds of cognitive testing, we narrowed down the self-reported items to three: one that assessed the number of days of medication missed in the last month, a second that assessed frequency of medication taking, and a third that asked patients to rate their medication taking. We then conducted a field test which showed that the three items had good internal consistency reliability [6]. In this paper we report the results of a study to assess the validity of these three self-report items, for both HIV antiretroviral medications, and also for non-HIV antiretroviral medications.

This validation study compares the self-report items to adherence data gathered using an electronic drug monitoring (EDM) device. We present analyses using two scoring approaches: one that assigns a score to each item response using a linear 0–100 scale (raw score), and a second that assigns a score to each item response based on the EDM result (calibrated score).

Methods

Participants and Study Procedures

Participants for this validity study were patients with HIV cared for at a hospital-based HIV specialty care practice. All study procedures were reviewed and approved by both the hospital and a university IRB. Enrollment criteria included being HIV-positive, currently being treated with an HIV antiretroviral medication and at least one other non-HIV related chronic medication, being over the age of 18 years, and having had a detectable viral load at one of the two most recent viral load assessments.

At the first or baseline visit, study participants completed a survey which included questions about sociodemographics, HIV risk factors, the three self-report adherence items, and other factors potentially related to medication adherence. At the baseline visit participants also were given a MedSignals® cellular medication monitoring device (http://

www.medsignals.com), and trained to use the device. This MedSignals device contains four bins or bays, allowing up to four medications to be monitored simultaneously. Each time the cover to a bay is opened, an event is registered by the software in the device, and these events are transmitted wirelessly to a server. Medications were chosen to be placed in each of the bays according to study protocols. All four bays were used, if possible.

The study design called for patients to make three additional study visits at one-month intervals after the baseline visit. At each of these subsequent visits, participants reported on their current prescription regimen, responded to the self-report items for each of the medications followed in the MedSignals bays, and responded to a subset of selected scales. This was a strictly observational study. There was no associated intervention. The reminder functions on the MedSignals device were disabled. Patients' physicians were not told of patients' adherence results.

Data Collection

Laptop computers were used to collect self-reports. For the self-report items pertaining to alcohol and drug use, demographics, and socioeconomic status, an ACASI approach was used. A CASI approach was used for medication adherence self-report items to minimize missing data.

We instructed patients in the use of a diary in which they could record times when they did not use the MedSignals device as instructed. For example, if a patient went out for an evening, or away for a weekend, and instead of taking the MedSignals device they took the pills they would need with them (that is, they "pocketed" the doses), then we asked them to record that in the diary. When the study ended, we collected the diaries. We also asked at each follow-up visit if the participant ever used the device in a non-standard way. When patients described events in their diaries or in the interview that were not captured by the MedSignals device, we inserted these presumed adherence events into the electronic record, which resulted in a total of 38 event changes for 22 participants.

Variables

Sociodemographic Variables

Sociodemographic variables included age in years, gender, educational level, race, ethnicity, HIV risk factor, sexual orientation, housing stability, income, primary language, and health insurance. Depression was assessed using the PHQ-9 [9], alcohol use was assessed using the AUDIT [10], and substance abuse was assessed using the Substance Abuse and Mental Illness Symptoms Screener (SAMISS) [11]. We used a PHQ-9 score of ten or greater as a cutoff for major depression, an AUDIT score of eight of greater as harmful, and a SAMISS score of greater than 9 as harmful.

Electronic Drug Monitoring (EDM) Adherence

Using the data from the MedSignals device, we calculated the percent of the time in an interval that was a participant was "covered" by medication [12–14]. Uncovered time began accumulating after a grace period following a missed dose that varied depending on the

dosing interval. We allowed 1, 2, and 3 h grace periods for 3-times a day, 2-times a day, and 1-time a day medications, respectively. Adherence was expressed as a percentage and calculated as the total minutes in the 30 days prior to a study visit minus the number of uncovered minutes in the same interval, divided by the total minutes in the interval ×100. We also calculated the percentage of doses taken, which correlated 0.90 with the covered time variable.

Self-report Adherence Variables

The three self-report items that we tested were as follows (see Appendix 1 for actual items):

1. *Days taken* In the last 30 days, on how many days did you miss at least one dose of any of your [drug name]? (write in number of days, 0–30). [Note that "days taken" is 30 minus the number of days missed.]

2. *Frequency* In the last 30 days, how often did you take your [drug name] in the way you were supposed to? (never/rarely/sometimes/usually/almost always/always)

3. *Rating* In the last 30 days, how good a job did you do at taking your [drug name] in the way you were supposed to? (very poor/poor/fair/good/very good/ excellent)

The rationale for these time frames, item wordings, and response options are described elsewhere [6, 7].

For analyses that used raw scores, item responses for the three adherence items were linearly transformed to a 0–100 scale with zero being the worst adherence, and 100 the best [15, 16]. Summary scales were calculated as the mean of the three individual items.

We calibrated each of the self-report items to the EDM response. We did this for two main reasons. First, each of the self-report items had a different response scale, and although we can put them all on a 0–100 scale, the assumption of linearity may not be correct. Second, a self-report measure is more useful to the extent that it approximates true or actual adherence. Calibration addresses both of these issues.

To implement the calibration, we first determined the average EDM adherence for each of the response options for each of the self-report options. For this part of the study we used responses to all study medications (HIV and non-HIV medications). When there were small number of responses in a category we aggregated adjacent categories (e.g., poor and very poor). We also aggregated categories if there was significant non-linearity in the EDM score for adjacent categories. We then took the mean of the three calibrated items to create a calibrated summary scale. We randomly selected 2/3 of the sample for this calibration (test set), and then tested the calibrated scores on the remaining 1/3 (validation set). Results from the test and validation sets were not different, so we recalculated the item scores on the whole sample and present these calibrated values in the results. The resulting calibration values are shown in Table 3 in Appendix 2.

Analyses

All analyses present data from HIV antiretrovirals and non-HIV antiretrovirals separately. We used descriptive statistics to show participants' sociodemographic characteristics, and rates of depression, alcohol use, and substance use. Individual self-report items and the three-item summary scale are described using means, medians, and ranges. We used Cronbach's alpha to demonstrate internal consistency reliability [17].

We used Pearson correlation coefficients (Spearman correlation coefficients were essentially identical) and difference scores to compare the self-report items and scales with the paired EDM measures using the covered time measurement derived from the MedSignals data. Time intervals for the self-reports and the EDM measures were exactly matched. To determine difference scores we subtracted the self-report score from the MedSignals score for each patient, and then calculated the mean difference score over all of the patients. We took this approach because the difference scores capture both the magnitude and direction of the relationship between the two scores. Because some patients have more than one ARV or non-ARV at a given visit, and because patients have up to three follow-up visits, adjustments for both of these types of clustering effects were needed. To accomplish this, we constructed generalized linear mixed models using SAS Proc Mixed to estimate the difference between the EDM and SR adherence. The model included 3 terms: measured by EDM or SR, HIV antiretroviral medication or non-HIV antiretroviral medication, and the interaction term (to estimate the mean for each of 4 categories with 2 terms). The model accounted for clustering within patients with specification of a normal distribution and the identity link.

We conducted similar analyses for the calibrated items and scales. In addition, we constructed receiver-operating characteristic (ROC) curves for the calibrated items and scales that used cut-offs of 0.80 and 0.90 by EDM to define adherent versus non-adherent patients.

Results

Follow-up Data and Participant Characteristics

Eighty-one participants completed the enrollment visit, and 71, 63, and 59 completed one, two, and three follow-up visits, respectively. Medication adherence self-reports were done at all four visits. MedSignals data were collected at the three follow-up visits. The numbers of HIV antiretroviral medications monitored at the three follow-up visits were 157, 138, and 136, respectively. The numbers of non-HIV antiretroviral medicines at the three follow-up visits were 74, 62, and 59, respectively.

The mean age of patients was 46 years, 37 % were female, 49 % had some education beyond high school, 22 % were Black, and 22 % were Hispanic. English was the primary language for 90 % of participants, and only 2 % reported no current place to live. Twenty-five percent had a score of ten or greater on the PHQ9 (moderate or greater depression), 15 % had an AUDIT score of eight or greater (problem drinking), and 38 % were positive on the seven-item substance abuse scale of the SAMISS (possible alcohol or substance abuse problem).

Monitored Medications

Of the 431 bin-periods of HIV ARVs, 82 were Emtricitabine/tenofovir (Truvada), 75 were Ritonavir (Norvir), 50 were Darunavir (Prezista), and 41 were Atazanavir (Reyataz). Of the non-HIV ARVs, 73 were mental health medications, 56 were antihypertensives, 18 were for elevated lipids, and 12 were for diabetes.

Descriptive Characteristics of Self-report and EDM Measures

Descriptive characteristics of the raw self-report items, 3-item scale, and EDM measure for HIV ARVs and non-ARVs are shown in Table 1. Mean scores for the days taken, frequency, rating items were 92.0, 84.0, 80.1, for HIV ARVs and 91.8, 83.2 and 80.0 for non-HIV ARVs, respectively. The percent of items at the ceiling, or 100, ranged from 48.8 to 56.3 for HIV ARVs and 47.5–54.9 for non-HIV ARVs. For the 3-item scale, the mean adherence scores were 85.1 and 84.7 for HIV ARVs and non-HIV ARVs, respectively, and the percent at the ceiling were 38.0 and 40.9, respectively.

For the EDM measure, mean adherence for HIV ARVs and non-HIV ARVs was 77 and 78, respectively, and the percent at the ceiling was 6.1 and 5.9 %, respectively.

Comparison Between Raw 3-Item Scale and EDM Measure

The raw 3-item scale and the EDM measure, for both HIV ARVs and non-ARVs, are compared in Fig. 1. The difference between the scores is 7.5 points (95 % CI 4.4–10.5, p <0.0001) and 5.2 points (95 % CI 0.7–9.7, p = 0.02) for the HIV ARVs and the non-ARV's, respectively. In both cases the self-report scale is statistically significantly higher than the EDM measure.

Correlations Between Self-Report and EDM Measures

Correlations between the self-report and the EDM measures are shown in Table 2. For the individual self-report items for HIV ARVs the correlations ranged from 0.27 to 0.43, and for non-ARVs they ranged from 0.38 to 0.62. The correlations for the two scales were 0.41 and 0.54, respectively. The Cronbach's alpha was 0.84 for the HIV ARVs and 0.87 for non-ARVs.

Calibrated Item and Scale Scores

Figure 2 shows the difference between the raw score and the calibrated scores for each of the 3 self-report items. The raw days taken item score is approximately parallel to the calibrated score, but is about 10 scale points higher. The frequency and rating items show a different pattern, with the raw and calibrated items having different slopes. In both cases, the raw and calibrated measures are approximately the same for the second to highest Likert scale category, "almost always" and "very good", respectively. Also, in both cases, the raw scale overestimates EDM adherence at the high end of the scale and underestimates it at the low end of the scale.

Figure 3 compares the raw scale, calibrated scale, and EDM measures, showing how the calibration worked to adjust or correct the overestimation of EDM adherence that the raw

scale demonstrated. The Pearson correlation coefficients between the calibrated scores and the EDM scores were 0.47 and 0.51 for HIV ARVs and non-ARVs, respectively.

Using a cut-off of 0.8 to define acceptable adherence, the c-statistics for the calibrated scale for HIV ARVs and non-ARVs were 0.75 and 0.76, respectively. Using a cutoff of 0.90, the c-statistics were 0.75 and 0.77, respectively.

Discussion

There are four main findings from these analyses. First, the three-item scale minimally overestimated the objective EDM adherence measure when using the raw score approach. Second, the pattern seen for HIV ARVs is similar to that seen for non-ARV medications. Third, a simple calibration process can improve the accuracy of the self-report scale by making it more closely approximate the objective EDM measure. Fourth, the calibrated measure showed good discrimination when using the EDM measure as a gold standard.

Several recent studies have described differences between self-reported and electronic drug monitoring devices. Thiramurthy et al. [18] studied Kenyan patients newly initiating antiretroviral therapy, and found a difference of 12.7 %. Shi et al. [19] reviewed the literature and found that self-reports were on average 14.9 % higher than electronic drug monitoring measures. They also conducted a meta-analysis which showed a pooled correlation coefficient of 0.45 [20]. Amico et al. conducted a detailed analysis of patients in the iPrEx study, using self-reported measures, pharmacy measures, and drug levels [21]. They found "large discrepancies" between self-report and pharmacy measures and drug levels. The overestimations that we found for the 3-item scale, 7.5 (7.45 with 2 decimal points) points for HIV ARVs and 5.2 points for non-ARVs, were smaller than those previously reported. There are probably two main reasons for these smaller differences. First, these items were developed using rigorous methods that included four rounds of iterative cognitive testing in a diverse sample of English-speaking participants. Likert-type response scales may work better than other response options because they probably map more clearly to the cognitive processes that patients use in practice when they formulate responses [6, 7, 22]. Second, the validation study was done in a setting in which there were no incentives for patients to inflate their responses, as there can sometimes be in clinical trials [21].

The cognitive testing described in previous publications [6, 7] found no differences in the cognitive processes that patients employed in responding to adherence items for HIV ARV and non-ARV medications, and the results of these validation tests suggest that the items perform nearly identically for these different medication classes. These findings suggest that the three-item scale can be used as a generic medication adherence measure.

Calibration is not commonly done with self-reported measures because there are rarely objective or "gold standard" measures with which to implement a calibration process. The calibration eliminated the upward bias usually seen with self-report measures. Note also that the calibration accomplished different things with the frequency and rating items that it accomplished with the days missed item. The days missed item is, in effect, parallel to the EDM score, so the calibration accomplishes a mean adjustment equal to the difference in the

Y-axis distance between the two self-report item curves. For the frequency and rating items, the slopes of the items are different from the slopes of the EDM curve, producing overestimation at higher levels of adherence, and underestimation at lower levels. The calibration addresses both the over- and under-estimations.

Although the days missed item consistently overestimates true or EDM adherence, it is the item that showed the highest correlation with the EDM adherence. One hypothesis for this stronger correlation is that there are aspects of this item stem and response task that make it easier for patients to recall. Work from the marketing literature shows that people are able to use enumeration or counting of events, rather than estimation methods, when there are fewer than five events during the recall period [22, 23]. In another study participants were more likely to use enumeration methods for irregular rather than regular behaviors, and for dissimilar rather than similar behaviors [24]. If missing a day of medications is a relatively uncommon event, and happens for reasons that are somehow salient or memorable, this might explain the higher correlation for the days missed item. Another factor may be that the missed days response option allows for finer detail in the response. The scale development process purposely explored both enumeration and estimation approaches, and the final scale includes one enumeration and two estimation items because different people probably use different combinations of these approaches [6, 7].

The discrimination of the calibrated 3-item scale, as assessed by the c-statistic, was good, with AUCs of 0.75–0.77 using two different cut-offs (0.8 and 0.9). These analyses use the EDM value as the gold standard, and the 3-item scale that is calibrated using the EDM, so in a sense there is also some over-fitting here also. However, the AUCs for the raw scale were all over 0.70, so this over-fitting is minimal. These data suggest that the discrimination of the self-report scale is good.

This analysis has several limitations. First, it is a medium-sized study using subjects from a single site. While the population was quite diverse, with 37 % women, only 49 % with education beyond high school, 22 % Black race, and 22 % Hispanic/Latino ethnicity, our results may not be generalizable to other dissimilar populations. Second, the approach we used to implementing the calibration is, in effect, a type of prediction model, in that we use self-report to predict the "true" EDM score. The results are likely to be over-fitted to the population that we used to develop the model [25], and the calibrations should be repeated in other sample populations. Third, we recognize that the assessment of adherence using pill container openings has many of its own limitations, including patients pocketing pills or otherwise not using the pill container as directed, and the fact that pill container opening is not a measure of actual pill ingestion [26]. Finally, we tested items developed with English-speaking populations. Translation into other languages, and subsequent examination of the performance characteristics of the self-report scale in those languages, needs to be carefully and rigorously done.

For purposes of this validation study, we asked patients about each medication that we monitored with the MedSignals device. However, when assessing adherence for patients using multi-pill ART regimens, we do not believe that it is necessary to ask about each pill. A number of studies have examined this issue, and all have concluded that differential

adherence is uncommon [27–30]. We conducted a randomized study which compared an item that asked about a single ARV to an item that asked about a patient's full ART regimen, and found that mean adherence levels were nearly identical [6]. Based on this evidence we recommend that both clinicians and researchers use a single item to measure ART adherence when patients are using multi-pill regimens.

Both clinicians and researchers sometimes want a measurement tool that will classify patients as adherent versus non-adherent. This 3-item scale was developed to be a continuous, not a dichotomous, measure. Because different medications have different pharmacokinetic and pharmacodynamic profiles that result in varying pharmacologic "forgiveness" [31], it would not be appropriate to specify any single cut off as denoting an "adherent" patient, either for HIV antiretrovirals [32], or for other medications [33].

In conclusion, using EDM as a gold standard, we conducted validity testing of a rigorously developed three-item medication adherence self-report scale, using both HIV antiretroviral medications and other medications. Our results support the validity of the self-report scale, and suggest that it can be used for all classes of medications. In clinical care the scale is probably best used as a screening tool that can identify people with non-perfect adherence who would be candidates for a more detailed, face-to-face assessment of potential barriers to optimal medication taking. In clinical research, or in quality improvement work that compares populations of patients, this self-report scale can be a useful and efficient tool when more complex and expensive methods such as EDM cannot be used, or when pharmacy refill data are unavailable.

Acknowledgments

References

1. Rand, CS. "I took the medicine like you told me, Doctor": self-reports of adherence with medical regimens. In: Stone, AA.; Turkkan, JS.; Bachrach, CA.; Jobe, JB.; Kurtzman, HS.; Cain, VS., editors. The science of self-report: implications for research and practice. Vol. 1. Mahwah: Lawrence Erlbaum Associates; 2000. p. 257-76.

2. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med. 2005; 353(5):487–97. [PubMed: 16079372]

3. Steiner JF, Prochazka AV. The assessment of refill compliance using pharmacy records: methods, validity, and applications. J Clin Epidemiol. 1997; 50(1):105–16. [PubMed: 9048695]

4. Demonceau J, Ruppar T, Kristanto P, et al. Identification and assessment of adherence-enhancing interventions in studies assessing medication adherence through electronically compiled drug dosing histories: a systematic literature review and meta-analysis. Drugs. 2013; 73(6):545–62. [PubMed: 23588595]

5. Kalichman SC, Amaral CM, Cherry C, et al. Monitoring medication adherence by unannounced pill counts conducted by telephone: reliability and criterion-related validity. HIV Clin Trials. 2008; 9(5): 298–308. [PubMed: 18977718]

6. Wilson IB, Fowler FJ Jr, Cosenza CA, et al. Cognitive and field testing of a new set of medication adherence self-report items for HIV care. AIDS Behav. 2013; 18(12):2349–58.

7. Fowler FJ, Lloyd SJ, Cosenza CA, Wilson IB. Coding cognitive interviews: an approach to enhancing the value of cognitive testing for survey question evaluation. Field Methods. 2014 1525822X14549921.

8. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care. 1986; 24(1):67–74. [PubMed: 3945130]

9. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA. 1999; 282(18):1737–44. [PubMed: 10568646]

10. Saunders JB, Aasland OG, Babor TF, De la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. Addiction. 1993; 88(6):791–804. [PubMed: 8329970]

11. Pence BW, Gaynes BN, Whetten K, Eron JJ Jr, Ryder RW, Miller WC. Validation of a brief screening instrument for substance abuse and mental illness in HIV-positive patients. J Acquir Immun Defic Syndr. 2005; 40(4):434–44.

12. Genberg BL, Wilson IB, Bangsberg DR, et al. Patterns of antiretroviral therapy adherence and impact on HIV RNA among patients in North America. AIDS. 2012; 26(11):1415–23. [PubMed: 22767342]

13. Genberg BL, Lee Y, Rogers WH, Willey C, Wilson IB. Stages of change for adherence to antiretroviral medications. AIDS Patient Care STDS. 2013; 27(10):567–72. [PubMed: 24093810]

14. Genberg BL, Lee Y, Rogers WH, Wilson IB. Four types of barriers to adherence of antiretroviral therapy are associated with decreased adherence over time. AIDS Behav. 2015; 19(1):85–92. [PubMed: 24748240]

15. Gaito J. Measurement scales and statistics: resurgence of an old misconception. Psychol Bull. 1980; 87(3):564–7.

16. Townsend JT, Hu GG, Evans RJ. Modeling feature perception in brief displays with evidence for positive interdependencies. Percept Psychophys. 1984; 36(1):35–49. [PubMed: 6493963]

17. Nunnally, JC.; Bernstein, IH. Psychometric Theory. New York: McGraw-Hill, Inc; 1994.

18. Thirumurthy H, Siripong N, Vreeman RC, et al. Differences between self-reported and electronically monitored adherence among patients receiving antiretroviral therapy in a resource-limited setting. AIDS. 2012; 26(18):2399–403. [PubMed: 22948266]

19. Shi LZ, Liu JN, Koleva Y, Fonseca V, Kalsekar A, Pawaskar M. Concordance of adherence measurement using self-reported adherence questionnaires and medication monitoring devices. Pharmacoeconomics. 2010; 28(12):1097–107. [PubMed: 21080735]

20. Shi L, Liu J, Fonseca V, Walker P, Kalsekar A, Pawaskar M. Correlation between adherence rates measured by MEMS and self-reported questionnaires: a meta-analysis. Health Qual Life Outcomes. 2010; 8(1):99. [PubMed: 20836888]

21. Amico KR, Marcus JL, McMahan V, et al. Study product adherence measurement in the iPrEx placebo-controlled trial: concordance with drug detection. J Acquir Immun Defic Syndr. 2014; 66(5):530–7.

22. Wilson IB, Carter AE, Berg KM. Improving the self-report of HIV antiretroviral medication adherence: is the glass half full or half empty? CurrHIV/AIDS Rep. 2009; 6(4):177–86.

23. Blair E, Burton S. Cognitive-processes used by survey respondents to answer behavioral frequency questions. J Consum Res. 1987; 14(2):280–8.

24. Menon G. The effects of accessibility of information in memory on judgments of behavioral frequencies. J Consum Res. 1993; 20(3):431–40.

25. Harrell FE, Lee KL, Mark DB. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996; 15:361–87. [PubMed: 8668867]

26. Wendel CS, Mohler MJ, Kroesen K, Ampel NM, Gifford AL, Coons SJ. Barriers to use of electronic adherence monitoring in an HIV clinic. Ann Pharmacother. 2001; 35(9):1010–5. [PubMed: 11573846]

27. Wilson IB, Tchetgen E, Spiegelman D. Patterns of adherence with antiretroviral medications: an examination of between-medication differences. J Acquir Immun Defic Syndr. 2001; 28(3):259–63.

28. McNabb JJ, Nicolau DP, Stoner JA, Ross J. Patterns of adherence to antiretroviral medications: the value of electronic monitoring. AIDS. 2003; 17(12):1763–7. [PubMed: 12891062]

29. Gardner EM, Burman WJ, Maravi ME, Davidson AJ. Selective drug taking during combination antiretroviral therapy in an unselected clinic population. J Acquir Immun Defic Syndr. 2005; 40(3):294–300.

30. Deschamps AE, De Geest S, Vandamme AM, Bobbaers H, Peetermans WE, van Wijngaerden E. Diagnostic value of different adherence measures using electronic monitoring and virologic failure as reference standards. AIDS Patient Care STDS. 2008; 22(9):735–43. [PubMed: 18754705]

31. Osterberg L, Urquhart J, Blaschke T. Understanding forgiveness: minding and mining the gaps between pharmacokinetics and therapeutics. Clin Pharmacol Ther. 2010; 88(4):457–9. [PubMed: 20856243]

32. Gulick RM. Adherence to antiretroviral therapy: how much is enough? Clin Infect Dis. 2006; 43(7):943–4.

33. Kravitz RL, Melnikow J. Medical adherence research: time for a change in direction? Med Care. 2004; 42(3):197–9. [PubMed: 15076818]

Appendix 1

In the last 30 days, on how many *days* did you miss at least one dose of any of your [drug name]?

Write in number of days: _____ (0–30)

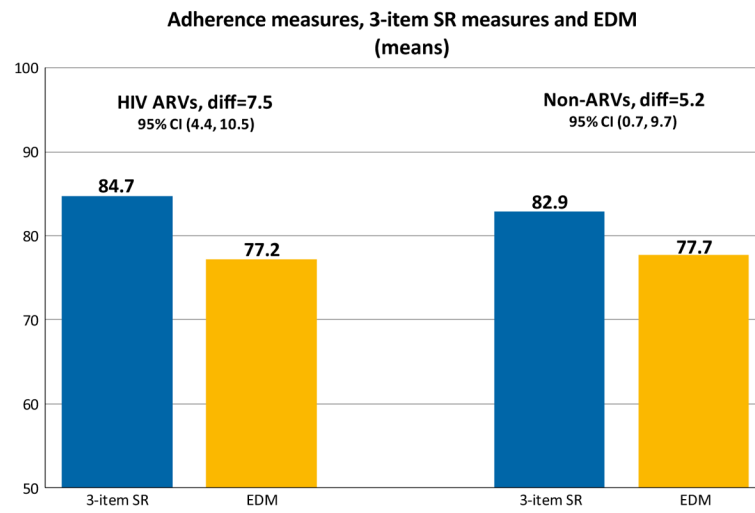In the last 30 days, how good a job did you do at taking your [drug name] in the way you were supposed to?

☐ Very poor

☐ Poor

☐ Fair

☐ Good

☐ Very good

☐ Excellent

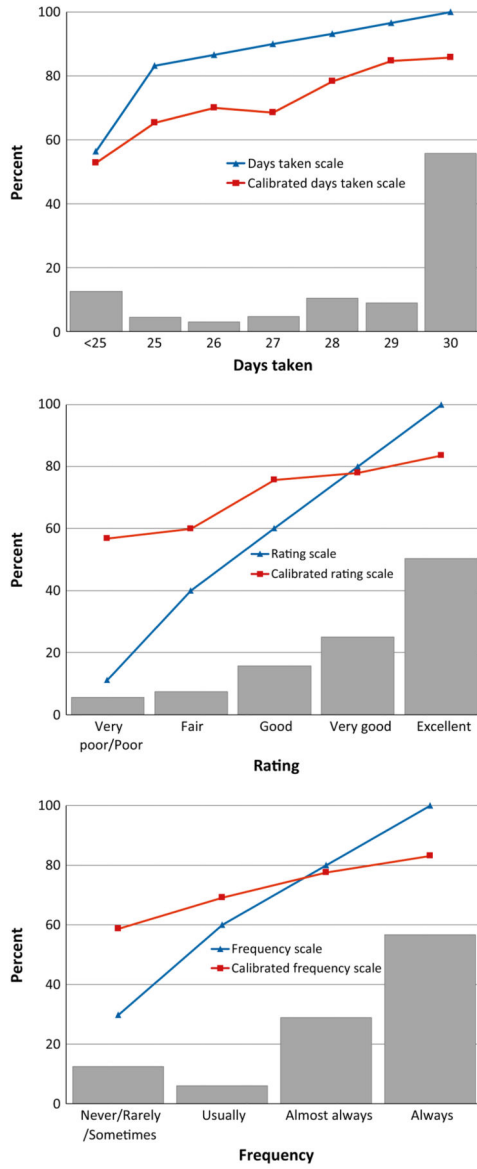In the last 30 days, how often did you take your [drug name] in the way you were supposed to?

☐ Never

☐ Rarely

☐ Sometimes

☐ Usually

☐ Almost always

☐ Always

Appendix 2

See Table 3.

Fig. 1.

Comparison of raw 3-item self-report scale with EDM measure for HIV ARVs (N = 431) and non-ARVs (N = 195). The 95 % CI's use mixed models to account for the clustering in the data structure

Fig. 2.
In this figure, this histogram shows the distribution of item scores for each of the 3 self-report adherence scales (including both HIV and non-HIV medications). The *blue line* shows the the score given to each response category using a linear, equal interval, zero to 100 approach, and the *red line* shows the score calibrated to the EDM measure (Color figure online)
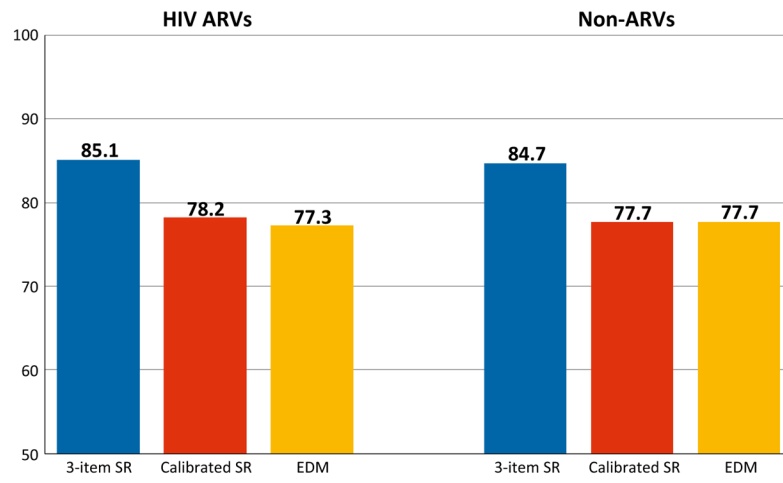
Fig. 3.
This figure shows mean scores for the raw scale, the calibrated scale, and the EDM measure for HIV ARVs and non-ARVs, respectively

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Descriptive characteristics of the raw self-report items, 3-item scale, and EDM measure for HIV ARVs and non-ARVs

Item/scale	N of bins	Mean (SD)	Median	Range	Percent at 100
HIV ARVs					
Days taken	581	92.00 (17.39)	100.00	0–100	56.28
Frequency	605	84.03 (23.75)	100.00	0–100	54.88
Rating	605	80.13 (25.53)	80.00	0–100	48.76
3-Item scale	605	85.05 (19.89)	93.33	0–100	38.02
EDM	377	77.25 (21.33)	84.23	21.78–100	6.10
Non-ARVs					
Days taken	264	91.82 (16.34)	100.00	0–100	54.92
Frequency	272	83.24 (23.08)	100.00	0–100	54.04
Rating	274	80.00 (24.09)	80.00	0–100	47.45
3-Item scale	274	84.65 (19.48)	91.11	13.33–100	40.88
EDM	170	77.69 (23.33)	87.41	13.19–100	5.88

Table 2

Correlations between the calibrated self-report and EDM measures for individual items and the three-item scale (all p <0.0001), and Cronbach's alphas

	All (N = 516)	HIV ARV (N = 354)	Non-ARVs (N = 162)
Days taken	0.50	0.43	0.62
Frequency	0.32	0.31	0.38
Rating	0.33	0.27	0.44
3-Item scale	0.45	0.41	0.54
Cronbach's alpha	–	0.84	0.87

Table 3

EDM values used to calibrate the self-report items

Levels of days taken	N of bins	Average EDM used for calibration
< 25	69	52.83
25	28	65.46
26	18	70.09
27	22	68.61
28	41	78.38
29	49	84.79
30	291	85.81

Levels of frequency	N of bins	Average EDM used for calibration
Sometimes or below	70	58.79
Usually	29	69.22
Almost always	159	77.51
Always	282	83.17

Rating	N of bins	Average EDM used for calibration
Poor or below	24	56.8
Fair	50	59.92
Good	71	75.65
Very good	143	77.87
Excellent	254	83.56