Published in final edited form as:

J Clin Epidemiol. 2011 March ; 64(3): 250–254. doi:10.1016/j.jclinepi.2010.07.014.

Improving the measurement of self-reported medication nonadherence

Corrine I. Voils^{a,b,*}, Rick H. Hoyle^c, Carolyn T. Thorpe^d, Matthew L. Maciejewski^{a,b}, and William S. Yancy Jr.^{a,b}

^aCenter for Health Services Research in Primary Care, Durham Veterans Affairs Medical Center, Durham, NC, USA

^bDepartment of Medicine, Duke University Medical Center, Durham, NC, USA

^cDepartment of Psychology and Neuroscience, Duke University, Durham, NC, USA

^dDepartment of Population Health Sciences, Health Innovation Program, University of Wisconsin, Madison, WI, USA

Abstract

Objective—Medication nonadherence is a significant clinical problem in chronic disease management. Self-report measures have inadequate reliability and poor distributional properties. We demonstrate how two fundamental measurement issues have limited the usefulness of self-reported medication nonadherence measures and offer recommendations for improving measurement.

Study Design and Methods—We reviewed existing self-report measures of medication nonadherence in the context of hypertension, one of the most common chronic conditions in which medication nonadherence is a paramount concern. We evaluated these measures with regard to two issues: (1) conflation of causal indicators (which give rise to a latent construct) and effect indicators (which are determined by a latent construct), and (2) a lack of evidence regarding the stability of nonadherence over time.

Results—Nonadherence measurement could be improved by using effect indicators to assess the extent of nonadherence and causal indicators to assess reasons for nonadherence. Moreover, nonadherence should be assessed longitudinally, so that recent developments in statistical modeling can illuminate the extent to which medication nonadherence is transient vs. stable.

Conclusion—Attention to these measurement issues can improve the assessment of self-reported nonadherence, thereby allowing more accurate conclusions to be drawn about medication-taking behavior and informing the development of improved interventions that target medication nonadherence.

Keywords

Adherence; Compliance; Psychometrics; Latent variable; Measurement; self-reports

1. Background

Medication nonadherence is associated with increased health care spending, high hospitalization and morbidity rates, and premature mortality [1,2]. Nonadherence averages 50% across different chronic diseases and medications [3].

Nonadherence may be measured directly or indirectly. Direct methods of assessing medication nonadherence detect the presence of the drug in a person's body using assays for the drug; its metabolites; or other markers in urine, blood, or other bodily fluids, but are used infrequently because of their cost and inability to provide feedback at the point of care. Indirect methods measure medication nonadherence behaviorally by means of electronic drug monitoring, pill counts, pharmacy refills, medical record review, directly observed therapy, clinician assessment, and self-reports. Although there is no “gold standard” for measuring nonadherence [1,4], the self-report method has unique advantages over other methods, including being brief, inexpensive, applicable in various settings, and able to provide immediate feedback at the point of care. Furthermore, self-reports can detect underlying issues contributing to nonadherence.

Despite these advantages, self-report measures are criticized for inadequate reliability and poor distributional properties (i.e., restricted range and skewness) [5]. Self-reported nonadherence rates are usually 10–20% lower than the rates obtained with other methods [4,6], leading to the conclusion that self-reported nonadherence does not reflect reality because of various biases (e.g., social desirability, inaccurate memory) [1,7]. We argue that some of the commonly cited disadvantages to self-reported nonadherence measures have arisen from inadequate attention to measurement issues.

The purpose of this commentary is to highlight two measurement issues contributing to poor measurement and to make recommendations regarding the ways of improving the reliability and validity of self-reported medication nonadherence measures. We selected a single therapeutic area, hypertension (HTN), in which the psychometric issues in existing nonadherence instruments could be illustrated. HTN is an ideal example because it is the most commonly coded chronic illness [8], most patients need multiple medications to control blood pressure, and treatment is generally taken lifelong [9]. Much effort has been devoted in assessing antihypertensive medication nonadherence, and multifactorial interventions have been developed to overcome it [10].

2. Existing self-report measures and their measurement issues

Self-report measures of antihypertensive medication nonadherence include single-item (e.g., how much medication one has missed [11]) and multiple-item measures. The most widely used multi-item measure is the four-item Morisky questionnaire, which inquires whether people miss doses because of forgetting, being careless, feeling better, or feeling worse [12]. Shea et al. modified the Morisky questionnaire by adding one question [13], and Morisky et al. recently expanded the scale to eight items [14]. These existing measures have addressed clinical and research needs to identify patients warranting intervention; however, measurement of nonadherence could be significantly improved by addressing the following measurement issues.

2.1. Measurement issue no. 1: conflating effect and causal indicator models of nonadherence

The first measurement issue concerns the conceptualization of medication nonadherence as a construct. As shown in Fig. 1, constructs may be measured directly (as a measured variable) or indirectly (as a latent variable). Measured directly, nonadherence could be operationalized

as the number or percentage of missed doses over a specified recall period. Direct measurement may be undesirable because it does not provide information on *why* people are not taking their medications as prescribed, which may be important for designing interventions.

Indirect measurement is often desirable because it can provide information on why people are not taking their medications. Indirect measurement refers to operationalizing a construct as a latent variable—that is, a variable—that is unobserved but is inferred from scores on questionnaire items (also known as indicators). For example, self-esteem is not directly observable but is inferred from responses to items, such as “I feel that I have a number of good qualities” and “On the whole, I am satisfied with myself” [15].

Once indirect measurement is chosen under the assumption that adherence is a latent construct, the investigator must choose between two measurement models that have divergent implications for measure development and validation: an effect indicator model or a causal indicator model. In an effect indicator model (Fig. 2A), the latent variable (e.g., self-esteem) determines the values of the indicators (e.g., items on a self-esteem measure), and item responses change as the latent construct changes. At least two, but usually more, highly intercorrelated questionnaire items are used to maximize the ability to observe the underlying latent construct. Many psychological constructs (e.g., self-esteem, caregiver burden, social support) have been measured properly with effect indicator models.

In contrast to the effect indicator model, the causal indicator model (Fig. 2B) involves indicators that are relatively unique, additive components of a construct. In this model, the omission of important components of a construct misrepresents it when measured. As a result, more than three or four items usually will be necessary to measure a construct adequately. An oft-used example of a construct that is measured with a causal indicator model is socioeconomic status (SES; the latent variable), which is determined by education, income, and occupation (causal indicators). A change in any indicator (e.g., pay raise) results in a change in the latent variable (SES). Causal indicators may be positively or negatively correlated, or uncorrelated, because they assess separate, distinct facets of a latent construct [16]. Sizable correlation between causal indicators is undesirable because each indicator should capture distinct aspects of the latent variable.

Existing multi-item measures of antihypertensive medication nonadherence—The Hill–Bone measure [17], the Brief Medication Questionnaire [18], and the Morisky and variations thereof [12–14]—contain items identifying respondents' *reasons* for nonadherence. The items cannot be appropriately classified as effect indicators for a couple of reasons. First, a person's level of nonadherence does not make them endorse an item describing one reason for nonadherence (e.g., forgetting); rather, forgetting to take a medication characterizes one's tendency toward nonadherence. Second, reasons for nonadherence may not be positively or highly correlated because the presence of one reason does not imply the presence of another. Nonetheless, low internal consistency is one of the oft-cited limitations of the Morisky measure [12], and one of the cited advantages of the newer eight-item measure [14] and the Hill–Bone Scale is the higher alpha coefficient [17].

These measures are more appropriately classified as causal indicator models. Yet, they do not represent all reasons for nonadherence, leading to distributional problems, such as restricted range and skewness, and prompting investigators to dichotomize responses (e.g., patients endorsing 1+ items are nonadherent) [18–20]. Dichotomization leads to information loss, misclassification of some individuals, and the floor effects and attenuated relationships with other nonadherence measures that have been widely observed.

The distinction between effect and causal indicator models also has two important implications for construct validation [16]. Internal consistency and factor analysis are appropriate because effect indicators should be positively and highly correlated. High internal consistency (Cronbach's alpha) depends on high interitem correlations, and factors from factor analysis account for covariance among items. These statistics are inappropriate for causal indicators, which may or may not be positively and highly intercorrelated. However, Cronbach's alpha is commonly used to assess the reliability of adherence measures that are structured (incorrectly) as causal indicator models.

2.2. Measurement issue no. 2: conceptualizing the stability of medication nonadherence over time

If nonadherence measures were appropriately construed as collections of causal or effect indicators, it would still be difficult to know how to intervene without knowing whether patients experienced a single episode of nonadherence (state-like nonadherence) vs. chronic nonadherence (trait-like nonadherence). Previous studies have been unable to differentiate these alternatives because adherence was assessed at a single time point. When assessing nonadherence longitudinally, investigators often vary the item wording (e.g., does one “ever” forget) or the recall period (e.g., 30 days instead of 7 days) rather than assessing the construct repeatedly over time. In a recent study, investigators concluded that nonadherence at one time point is a different construct than nonadherence over time because the proportions of doses taken over the previous 1–7 days were weakly correlated with a measure of general adherence tendencies [21].

Cognitive processes make it difficult to interpret nonadherence data obtained at a single time point because some people think of the specified recall period, whereas others think about their behavior patterns. Moreover, people may be reporting a socially desirable response or how they aspire to take their medications. A more accurate picture of what happens over time may be captured by people reporting their behaviors at several time points.

3. Recommendations for improving the measurement of nonadherence

Causal and effect indicator models have advantages in measuring medication nonadherence and in informing intervention design. An effect indicator–based measure could be used to detect the presence and extent of nonadherence, thereby identifying patients warranting intervention. A causal indicator model could be used to identify the complete set of barriers contributing to nonadherence, thereby informing the strategies used to increase adherence among targeted patients. Existing measures of medication nonadherence contain only effect indicators [11] or only causal indicators [12] or combine effect and causal indicators into a single score [13,18] without using the unique information provided by each.

A comprehensive approach incorporating both effect and causal indicator models may improve nonadherence assessment. Two-step approaches of identifying behavior problems in the clinic are not new and have proven useful [22]. For example, patients report on their experience of two common symptoms of depression (sad/blue and anhedonia) [22] to screen for depression. A positive response to either question prompts more extensive probing using a more extensive depression measure [23] or an in-depth interview by the provider. In the case of medication nonadherence, effect indicators could be used to assess the extent of nonadherence, after which positive responses could be followed by causal indicators to assess reasons for nonadherence.

Investigators should be cognizant of these two types of measures so that appropriate validity and reliability analyses will be performed when creating new measures or using existing measures. When creating new measures, researchers should specify the model a priori, so

that appropriate items can be generated and appropriate psychometric analyses can be conducted. When using existing measures, researchers should determine which type of model is implied by the items before proceeding with analyses. One method is to invoke “mental experiments,” whereby one imagines whether changes in responses to the items will cause changes in the construct or vice versa [24]. When ambiguous results are obtained, one may determine whether the model is based on causal or effect indicators by means of a vanishing tetrad test, which evaluates whether differences between pairwise covariances among indicators approach zero, denoting an effect indicator model [25]. If researchers use an effect indicator model, the traditional methods for assessing reliability and validity, including internal consistency and/or factor analysis, can be applied.

If researchers use a causal indicator model, alternative approaches for assessing reliability and validity should be considered. We recognize that reviewers may expect internal consistency estimates and/or factor analysis for any instrument. When using a causal indicator-based instrument, investigators should politely decline such requests and refer existing methods literature [16,26]. The central concern for validity assessment of causal indicator-based models is whether the items adequately represent all components of a construct. Expert input, literature review, and qualitative research studies could be used to generate an initial item pool. Researchers must then determine to what extent each indicator contributes to the ability of the measure to predict outcomes (e.g., does adding an indicator increase the prediction of criterion variables by adding to rather than duplicating other indicators?).

Researchers also need to use alternative methods for assessing the reliability of causal indicator-based measures. Test-retest reliability is appropriate for both types of measures. For effect indicators, one typically examines test-retest reliability of the total scale score; examining test-retest reliability of individual items is not necessary because items are positively, highly correlated. In contrast, for causal indicators, one must examine test-retest reliability of each item because the same total score could be obtained from different profiles of item responses, reflecting the fact that items do not have to be highly correlated. In such cases, a high test-retest correlation would be misleading. The challenge in assessing test-retest reliability in the context of nonadherence is determining the appropriate time interval that should separate the two assessments. Shorter intervals increase the risk of a learning effect or recall bias, which inflates reliability estimates. Longer intervals increase the risk for actual changes in the construct, which attenuates reliability estimates. Assuming nonadherence to be trait like (i.e., relatively stable over time), there is some leeway for this time period (e.g., 5–10 days). Assuming nonadherence to be state like (i.e., transient), the appropriate time period is more difficult to determine because low test-retest reliability may be reflective of true changes in the construct, unreliability, or both.

This challenge highlights the need for more evidence regarding the extent to which medication nonadherence is trait like vs. state like. Longitudinal evidence is needed to illuminate whether episodic or chronic nonadherence is more common. Advances in sophisticated longitudinal statistical methods allow such questions to be answered by separating variance in scores over time into stable (trait), unstable (state), and unreliability (error) components [27]. Trait-state-error models can be univariate or bivariate and can include time-varying (e.g., stress level) and time-invariant (e.g., conscientiousness) covariates that may predict patterns of medication-taking behavior. The latent state-trait model can be extended to determine whether there are subgroups of individuals (termed *latent classes*) who experience more change over time than others [28].

Considerable effort has been made trying to measure and to decrease medication nonadherence. To gain an accurate picture of people's medication-taking behavior, more

attention is needed on fundamental measurement issues, including the distinction between causal and effect indicator models and the determination of whether problems with adherence are chronic or episodic. Only then can interventions be appropriately designed and delivered. This article is a call for such efforts.

Acknowledgments

This research was supported by a grant from the National Institute of Aging (R21AG035233), by a Research Career Scientist award (RCS 10-391; to M.L.M.), and by resources and facilities at the Veterans Affairs Medical Center in Durham, NC. Views expressed in this article are those of the authors and do not necessarily represent the Department of Veterans Affairs.

References

- [1]. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. 2005; 353:487–97. [PubMed: 16079372]
- [2]. Sherbourne CD, Hays RD, Ordway L, DiMatteo MR, Kravitz RL. Antecedents of adherence to medical recommendations: results from the medical outcomes study. *J Behav Med*. 1992; 15:447–68. [PubMed: 1447757]
- [3]. Horne, R. Adherence to medication: a review of existing research. In: Myers, LB.; Midence, K., editors. *Adherence to treatment in medical conditions*. Harwood Academic Publishers; Amsterdam, The Netherlands: 1998. p. 285-310.
- [4]. Simoni JM, Kurth AE, Pearson CR, Pantalone DW, Merrill JO, Frick PA. Self-report measures of antiretroviral therapy adherence: a review with recommendations for HIV research and clinical management. *AIDS Behav*. 2006; 10:227–45. [PubMed: 16783535]
- [5]. Bangsberg D. Monitoring adherence to HIV antiretroviral therapy in routine clinical practice: the past, the present, and the future. *AIDS Behav*. 2006; 10:249–51. [PubMed: 16783537]
- [6]. Ingersoll KS, Cohen J. The impact of medication regimen factors on adherence to chronic treatment: a review of literature. *J Behav Med*. 2008; 31:213–24. [PubMed: 18202907]
- [7]. Bosworth, HB. Medication treatment adherence. In: Bosworth, HB.; Oddone, E.; Weinberger, M., editors. *Patient treatment adherence*. Lawrence Erlbaum Associates; Mahwah, NJ: 2006. p. 147-94.
- [8]. Cherry DK, Hing E, Woodwell DA, Rechtsteiner EA. National ambulatory medical care survey: 2006 summary. *Natl Health Stat Report*. 2008; (3):1–39. [PubMed: 18972720]
- [9]. Rosamond W, Flegal K, Furie K, Go A, Greenlund K, Haase N, et al. Heart disease and stroke statistics—2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2008; 117:e25–146. [PubMed: 18086926]
- [10]. Bosworth HB, Olsen MK, Neary A, Orr M, Grubber J, Svetkey L, et al. Take Control of Your Blood Pressure (TCYB) Study: a multifactorial tailored behavioral and educational intervention for achieving blood pressure control. *Patient Educ Couns*. 2008; 70:338–47. [PubMed: 18164894]
- [11]. Schroeder K, Fahey T, Hay AD, Montgomery A, Peters TJ. Adherence to antihypertensive medication assessed by self-report was associated with electronic monitoring compliance. *J Clin Epidemiol*. 2006; 59:650–1. [PubMed: 16713529]
- [12]. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care*. 1986; 24:67–74. [PubMed: 3945130]
- [13]. Shea S, Misra D, Ehrlich MH, Field L, Francis CK. Correlates of nonadherence to hypertension treatment in an inner-city minority population. *Am J Public Health*. 1992; 82:1607–12. [PubMed: 1456334]
- [14]. Krousel-Wood M, Islam T, Webber LS, Re RN, Morisky DE, Muntner P. New medication adherence scale versus pharmacy fill rates in seniors with hypertension. *Am J Manag Care*. 2009; 15:59–66. [PubMed: 19146365]
- [15]. Rosenberg, M. *Society and the adolescent self-image*. Princeton University Press; Princeton, NJ: 1965.

- [16]. Bollen K, Lennox R. Conventional wisdom on measurement: a structural equation perspective. *Psychol Bull.* 1991; 110:305–14.
- [17]. Kim MT, Hill MN, Bone LR, Levine DM. Development and testing of the Hill-Bone compliance to High Blood Pressure Therapy Scale. *Prog Cardiovasc Nurs.* 2000; 15:90–6. [PubMed: 10951950]
- [18]. Svarstad BL, Chewning BA, Sleath BL, Claesson C. The Brief Medication Questionnaire: a tool for screening patient adherence and barriers to adherence. *Patient Educ Couns.* 1999; 37:113–24. [PubMed: 14528539]
- [19]. Trivedi RB, Ayotte B, Edelman D, Bosworth HB. The association of emotional well-being and marital status with treatment adherence among patients with hypertension. *J Behav Med.* 2008; 31:489–97. [PubMed: 18780175]
- [20]. Prado JC Jr, Kupek E, Mion D Jr. Validity of four indirect methods to measure adherence in primary care hypertensives. *J Hum Hypertens.* 2007; 21:579–84. [PubMed: 17443212]
- [21]. Jerant A, DiMatteo R, Arnsten J, Moore-Hill M, Franks P. Self-report adherence measures in chronic illness: retest reliability and predictive validity. *Med Care.* 2008; 46:1134–9. [PubMed: 18953223]
- [22]. Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV 3rd, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA.* 1994; 272:1749–56. [PubMed: 7966923]
- [23]. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. *JAMA.* 1999; 282:1737–44. [PubMed: 10568646]
- [24]. Bollen, K. Structural equations with latent variables. Wiley; New York, NY: 1989.
- [25]. Bollen K, Ting K. A tetrad test for causal indicators. *Psychol Methods.* 2000; 5:3–22. [PubMed: 10937320]
- [26]. MacCallum RC, Browne MW. The use of causal indicators in covariance structure models: some practical issues. *Psychol Bull.* 1993; 114:533–41. [PubMed: 8272469]
- [27]. Kenny D, Zautra A. The trait-state-error model for multiwave data. *J Consult Clin Psychol.* 1995; 63:52–9. [PubMed: 7896990]
- [28]. Courvoisier D, Eid M, Nussbeck F. Mixture distribution latent state-trait analysis: basic ideas and applications. *Psychol Methods.* 2007; 12:80–104. [PubMed: 17402813]

What is new?

- Existing self-report measures of nonadherence have inadequate reliability and poor distributional properties.
- Existing measures do not inform the extent to which nonadherence is stable vs. unstable over time.
- Measurement of nonadherence could be improved by (1) attending to the distinction between effect and causal indicators and (2) assessing nonadherence longitudinally.

Fig. 1.
Decision tree: measuring a construct.

Fig. 2. Path diagrams of effect indicator (A) and causal indicator (B) measurement models. e , measurement error; d , disturbance or measurement error associated with prediction.