



Published in final edited form as:

HIV Clin Trials. 2012 ; 13(5): 271–277. doi:10.1310/hct1305-271.

Feasibility and Reliability of Interactive Voice Response Assessment of HIV Medication Adherence: Research and Clinical Implications

Jennifer E. Hettema, Sharzad Hosseinbor, and Karen S. Ingersoll

Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, Virginia, USA

Abstract

Background—There are well-documented negative consequences of nonadherence to HIV medications. Telephone-based interactive voice response (IVR) technologies may hold promise for assessing nonadherence in both research and clinical contexts; however, little psychometric research has been conducted on this topic.

Objective—In the present pilot study, we test the feasibility and reliability of a simplified patient-initiated, daily IVR system with a convenience sample of HIV patients attending a university-affiliated infectious disease clinic.

Methods—Participants were asked to call in to an IVR system to report adherence daily during 2 weeks of a larger prospective study. Response rates and patterns were analyzed for feasibility and compared to retrospective, self-report timeline follow-back (TLFB) adherence reporting.

Results—The IVR protocol showed moderate feasibility, with participants reporting adherence behavior on 63.4% of days. However, agreement with TLFB data was low, particularly for days in which participants reported incomplete adherence.

Conclusions—The IVR protocol tested in the current trial shows some promise. Completion rates were higher than in previous trials. Future research is needed to further enhance the feasibility of IVR for HIV medication adherence and to compare responses to more objective measures on HIV adherence.

Keywords

adherence, HIV, interactive voice response, measurement

Adherence to HIV treatment medications has significant individual and public health consequences.¹ Although nearly perfect levels of adherence result in reduced viremia and symptoms, even episodic nonadherence can lead to medication-resistant HIV.² Accurately measuring adherence and identifying interruptions in real time are necessary elements for clinicians to effectively intervene with patients, as viral rebound may occur following as little as a few days of nonadherence.³ Similarly, valid adherence outcome measures are crucial for researchers interested in measuring the virological and medical effects of adherence or in testing the efficacy of interventions to increase adherence. Telephone and mobile technologies are a promising means of assessing antiretroviral adherence, because of

their widespread use, ability to reach rural and marginalized populations, convenience, potential to assess adherence prospectively, and related potential for triggering timely intervention in real time before viral rebound. Such technologies may also have potential utility for promotion of adherence through 1-way or 2-way reminders or intervention systems.⁴ Traditional clinical and research tools for detecting nonadherence rely mainly on self-report or objective assessment (medication event monitoring system [MEMS] cap, pharmacy records, changes in laboratory values) that can only be assessed retrospectively or during points of patient contact. Some have argued that telephone technologies could induce a paradigm shift in health care delivery and the collection of patient information,⁵ but little is known about the reliability of telephone-based HIV adherence assessment.

Interactive voice response (IVR) is a telephone-based technology that uses prerecorded voice prompts to gather information.⁶ IVR has been used to collect health information for a variety of health behaviors, including safe sex practices⁷ and alcohol use⁸ in HIV prevention research. Although IVR and other mobile technologies are increasingly being used to assess behavior in HIV clinical and research settings,^{9,10} the relationship between these data collection methods and more traditional procedures is unknown. In a small pilot study (n=13) to determine the feasibility and acceptability of IVR to collect adherence data from caregivers of HIV-infected children, only 12.5% of the weekly IVR assessment calls resulted in successful responses.¹¹ In that study, participants received weekly researcher-initiated IVR assessment calls during which they were asked to enter their participant ID on a key pad and then key in the number of missed medication doses in the past 7 days, 30 days, or time since last refill. Qualitative interviews with participants revealed several important barriers to IVR response including conflicting needs for phone at time of call (taking other calls, sharing phone with others), poor reception, dead batteries, an inability to correct mistaken key entry, and confusion about entry of participant ID numbers. In the present pilot study, we investigate the feasibility and reliability of a simplified patient-initiated, daily IVR system that allows for voice versus key pad reporting.

METHOD

Participants

A convenience sample of HIV-positive individuals attending a university-affiliated infectious diseases clinic volunteered to participate in several study assessments as part of a larger study of the relationships between antiretroviral medication adherence and other patient characteristics and behaviors. Sample size for this substudy was based on a power analysis for the larger clinical study.

Procedures

Participants were recruited via flyers and brochures in clinic waiting rooms and referral by clinic providers to study staff. As part of the parent study, participants completed a baseline assessment that gathered detailed demographic information, self-reports of engagement in healthy and risky behaviors, and prescribed regimen, including names of all prescribed medications, pills per dose, and doses per day. During weeks 2 and 4 post baseline, participants were asked to call in daily to the IVR system, which was essentially a voicemail line that queried participants to prospectively report their participant ID and the number of medications taken on the previous day. One month after the baseline, participants recalled their medication adherence for the past month using the established timeline follow-back (TLFB) procedure. The feasibility of the IVR system and the relationship between IVR and traditional self-report measures of medication adherence during weeks 2 and 4 of the 1-month period are the focus of the current investigation.

Measures

The IVR was a standard answering machine on which participants verbally reported their participant number and the number of doses of each medication taken on the previous day. For example, participants were trained to state, “Hi, this is Joe Smith, and yesterday I took both doses of my Combivir.” Daily adherence was calculated as the number of pills taken on a particular day divided by the number of pills prescribed to be taken (as calculated by study staff from the patient's current prescribed regimen at baseline). Mean adherence for each participant was also calculated by averaging the daily adherence reports for each participant.

TLFB¹² was used to retrospectively measure medication-taking behavior for the past 30 days. The TLFB has been used extensively to measure substance use and other health-related behaviors.^{13,14} Several HAART adherence treatment outcome studies have used the TLFB modified for HIV medication adherence.^{15–18} The TLFB uses visual aids to guide collection of self-reported information on each participant's medication adherence. Staff members asked participants to recall HAART medication-taking behavior for each day over the previous month, using a guided method, pictures of their medications, and a calendar to prompt recall. This technique results in accurate recollection of daily behavior in clinical populations with excellent validity and reliability.^{12,13} TLFB also yields daily and mean adherence estimates based on proportion of pills per day.

Inter-instrument Reliability

IVR and TLFB adherence reports were compared for weeks 2 and 4 of the 1-month reporting period to determine inter-instrument reliability. Daily adherence (the proportion of pills taken on a specific day for which both data types are available) estimates for IVR and TLFB were compared. Inter-instrument reliability was calculated using the concordance correction coefficient (CCC),^{19,20} a type of intraclass correlation coefficient (ICC) that does not require assumptions of normality, because it is based on differences between observations made on the same subject. Previous research on adherence suggests that adherence data are typically not normally distributed but highly negatively skewed, with the majority of reporting days reflecting perfect adherence.²¹ Overall, ICCs provide a more conservative estimate of reliability than Pearson correlations, as they consider the absolute differences (vs relative rank orderings) between pairs of scores and thus control for the amount of similarity that may be a product of chance alone.²² For continuous variables, McBride²³ suggests the following interpretation of CCC: $\rho_c > .99$, almost perfect; $\rho_c = .95-.99$, substance; $\rho_c = .90-.95$, moderate; $\rho_c < .90$, poor. Agreement between continuous estimates of daily adherence was also calculated using a Bland Altman plot.^{24–27} By plotting the average of the 2 measurement techniques against the difference of each pair, this method produces an estimate of bias that quantifies the differences between measurement techniques. These differences can then be interpreted based on their clinical significance. The MedCalc software program (MedCalc; Mariakerke, Belgium; <http://www.medcalc.be/>) was used to calculate CCCs and create the Bland Altman plot.

In addition to analyzing data in a continuous fashion, adherence data were also transformed into categorical data, reflecting daily and mean perfect (100%) or imperfect (<100%) adherence. With these categorical data, kappa coefficients were used to estimate the proportion of agreement among raters after chance agreement has been removed. Kappa data distribution assumptions are less restrictive than those for ICC and include independence of raters and subjects and mutually exclusive, exhaustive rating categories.²⁸ Kappa values can range from -1 to $+1$, with a score of 0 indicating that agreement is not better than chance, a score of 1 indicating perfect agreement, and a score of -1 indicating perfect disagreement. Landis and Koch²⁹ suggest the following interpretation of positive kappa values: $<.40$, poor; $.40$ to $.75$, fair to good; $>.75$, excellent agreement.

RESULTS

Participant Characteristics

A total of 63 individuals participated in the investigation. A majority of participants were male (64%), African American (86%), and unemployed (83%), while 50% were heterosexual. On average, patients were 42.8 (*SD* 7.5) years of age. Participants were prescribed a mean of 6.6 (*SD* 4.2) pills per day and had been taking HIV medications for 53 months (*SD* 46). At the clinical lab testing nearest to the start of the study, participants had a mean CD4 count of 379 (*SD* 289) and a viral load of 51,137 (*SD* 144,360), indicating moderate disease progression.

IVR Feasibility

Most participants ($n=52$; 82.5%) called in to use the IVR on at least 1 occasion. As a whole, participants called in to report medication adherence a mean of 8.87 (*SD* 5.62) out of the 14 assigned days (63.4%). Among those who called in at least once, reporting rates were high, with participants phoning in to report adherence data on 10.55 (*SD* 4.43) out of the 14 days (75.4%).

IVR Reporting START

Of those who called in to IVR at least once, the mean self-reported total proportion of pills taken was high (mean = .96, *SD* .10; median = 1.00; inter-quartile range [IQR], 0.00) with 80.8% (42/52) of participants reporting perfect adherence on all call-in days. Imperfect mean adherence ($n=10$ participants) ranged from 0.56 to 0.98 with a mean of 0.81 (*SD* 0.15; median = 0.88; IQR, 0.28). Among all participants, 549 IVR daily adherence reports were provided. Mean daily adherence was 0.92 (*SD* .14; median = 1; IQR, 0.00). Perfect daily adherence was reported for 90.35% (496/549) of days. The IVR never received a daily report of complete noncompliance (0% of pills taken); however some participants reported taking less than their prescribed dose. Imperfect daily adherence ranged from 0.44 to 0.88 with a mean of 0.60 (*SD* 0.13; median = .50; IQR, .14).

Inter-instrument Reliability

Corroborating TLFB data was available for 74.6% ($n=47$) of the sample, consistent with follow-up rates in self-report research in similar populations.¹⁵ Mean adherence data were available from both measurement instruments for 40 participants (63.5%). Of the 14 days for which overlapping IVR and TLFB daily adherence data were possible, both data sources were available for 73.4% (411/560) of days.

As expected, visual inspection revealed that IVR and TLFB data were not normally distributed and were highly negatively skewed, with the majority of reporting days showing perfect adherence. A CCC comparing adherence reports for the days both data sources were available suggests poor agreement between IVR and TLFB methods ($\rho_c=.78$; 95% CI, .74 to .81). Supporting these findings, creation of a Bland Altman plot revealed that the average discrepancy between methods (bias) was .14 (*SD* 2.84; 95% CI, -.14 to .42), which is clinically significant and suggests poor agreement between measures.

To further understand the nature of disagreement, total (the mean proportion of pills taken across all reporting days for which both data types were available) and daily (the proportion of pills taken on a specific day for which both data types are available) data were transformed to categorical variables to prepare for kappa analysis. Adherence was defined as perfect or imperfect. Additional categories of imperfect adherence were not created because of the limited number and range of imperfect IVR adherence reporting days. Figure 1 is a graphical display of agreement between the 2 measures on total adherence. There was

agreement of perfect mean adherence in 29 of the 40 cases (72.50%) with data available from both sources. However, there were no instances of agreement regarding imperfect mean adherence. In the case of disagreements, each instrument appears to be equally likely to underestimate rates of adherence than the other, with TLFB estimates exceeding IVR estimates for 6 cases and IVR estimates exceeding TLFB estimates in 5 cases. Despite high levels of agreement about perfect mean adherence, kappa estimates of agreement were low ($\kappa = .02$; 95% CI, $-.30$ to $.34$).

Analysis of daily adherence data reveals a similar pattern. Agreement between the 2 reporting methods regarding perfect adherence occurred on 83.94% of the days (345/411 days), and there were no instances of agreement about imperfect adherence. TLFB estimates exceeded IVR estimates on 44 (10.71%) days and IVR estimates exceeded TLFB estimates on 22 (5.35%) days. Kappa estimates of the agreement about daily adherence by the 2 measures were also low ($\kappa = -.08$; 95% CI, $-.10$ to $-.06$).

DISCUSSION

In this pilot study, daily participant-initiated IVR showed moderate levels of feasibility. While rates of successful reporting were higher than those reported in a previous researcher-initiated weekly IVR study (63.4% in our study vs 12.5% in Haberer et al¹¹), they are still lower than acceptable rates needed for rigorous trials, limiting their research applications. Additionally, gaps in reporting days may limit the utility of IVR for triggering interventions that target nonadherence in clinical settings. Though reactivity to IVR was not assessed in this study, it is possible that such a system may promote improved adherence, as daily self-monitoring alone has been shown to improve medication adherence,³⁰ and such potential effects should be assessed in future research.

The improved rates of successful reporting that we observed compared to that observed in Haberer et al¹¹ suggest that specific aspects of the current IVR procedure may have improved compliance with IVR. We suggest that 2 key modifications that likely improved system compliance in the case of the current study were (1) simplification of the system by allowing voice reporting, rather than requiring keypad reporting, and (2) the flexibility in time and device used to report adherence resulting from patient versus researcher initiation.

While IVR and telephone-based adherence reporting have many potential benefits beyond traditional self-report or other objective measures, rates of interrater agreement between IVR and TLFB self-report are clearly suboptimal. The current study demonstrated that simplified IVR is feasible, but additional work is needed to identify factors that further increase response rate and to compare reported adherence to a non-self-report objective standard. Even though some researchers have found that self-report is likely to overestimate adherence compared to more objective measures, in general there is a high correlation between the two.³¹

Several methodological limitations should be considered when interpreting the results of the current study. First, the sample was composed of a convenience sample of patients who volunteered to participate in the study. Even though the investigator made attempts to recruit participants during a range of clinic days and hours, it is possible that the sample is not representative of all patients attending the clinic. A volunteer bias may have also created an unrepresentative sample. Additionally, even if the sample is representative of the clinic population, this population may not be representative of the larger HIV population, particularly those from international or resource-limited settings, potentially limiting the generalizability of findings. Moreover, participants knew the schedule of days on which they were to call in about their adherence, and this may have influenced their reporting. Research

on medication adherence in general, and antiretroviral medication adherence specifically, has been criticized for its lack of methodological rigor and failure to use validated measurement techniques.³² Several improvements have been recommended to improve the validity of adherence measurement and classification systems, including nonscheduled measurement,³² combining multiple measures,³³ using validated self-report measures,³² ensuring the confidentiality of participants,² and utilizing short recall periods.³⁴ IVR is consistent with all of these empirically grounded methodological recommendations. Although daily reporting would not likely be feasible for assessors or patients in the traditional in-person modality, the telephone modality has low time and cost burdens to both patients and providers and may reduce recall errors by using a very short recall period, asking for recall of only 1 day. However, it should be noted that daily reporting of medication adherence may be a burden to patients, especially those with a lifelong illness requiring chronic treatment. Currently, we view IVR as a potential enhancement to adherence assessment in a research context only. However, it is certainly possible to foresee telephone apps that could facilitate such reporting even for patient populations in the future, as the population becomes saturated with smart mobile phones.

This study demonstrates the need for further modification of IVR to enhance feasibility and further evaluation of the psychometric characteristics of IVR methods to gather data on HAART adherence. Further evaluation of the agreement of these methods with other measures of antiretroviral adherence such as electronic monitoring or plasma monitoring seems warranted. Although we are optimistic about the potential of IVR to gather reliable adherence data, we recommend against relying on IVR as a primary outcome measure of adherence in studies and suggest caution in applying this method in a clinical setting, until the field improves in methodology and reliability.

Acknowledgments

Funding support: This study was funded by K01 MH01688 from the National Institute of Mental Health to Dr. Ingersoll. Additional investigator effort was supported by R01 DA016554 from the National Institute of Drug Abuse.

Additional contributions: We thank the staff of the Virginia Commonwealth University Medical Center Infectious Disease Clinic for referring participants and thank the patients who volunteered for the study. Additionally, we thank research assistants Jessye Cohen, Arica Fishback, Kris Hash, Raphael Mutepa, Christina St. Clair, and Carrie Walker for data collection and management.

REFERENCES

1. UNAIDS. Global HIV/AIDS Response: Epidemic Update and Health Sector Progress Towards Universal Access. World Health Organization; Geneva: 2011.
2. Nieuwkerk PT, Oort FJ. Self-reported adherence to antiretroviral therapy for HIV-1 infection and virologic treatment response: a meta-analysis. *J Acquir Immune Defic Syndr*. 2005; 38:445–448. [PubMed: 15764962]
3. Parienti JJ, Das- Douglas M, Massari V, Guzman D, Verdon R, Bangsberg DR. Not all missed doses are the same: sustained NNRTI treatment interruptions predict HIV rebound to low-to-moderate adherence levels. *PLoS One*. 2008; 3:e2783. [PubMed: 18665246]
4. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. *Cochrane Database Syst Rev*. 2012; 14:CD009756. [PubMed: 22419345]
5. Shet A, de Costa A. India calling: harnessing the promise of mobile phones for HIV healthcare. *Trop Med Int Health*. 2011; 16:214–216. [PubMed: 21371214]
6. Schroder KEE, Johnson CJ. Interactive voice response technology to measure HIV related behavior. *Curr HIV/AIDS Rep*. 2009; 6:210–216. [PubMed: 19849964]

7. Schroder KEE, Johnson CJ, Wiebe JS. Interactive voice response technology applied to sexual behavior self-reports: a comparison of three methods. *AIDS Behav.* 2007; 11:313–323. [PubMed: 16841191]
8. Barta WD, Portnoy DB, Kiene SM, et al. A daily process investigation of alcohol-involved sexual risk behavior among economically disadvantaged problem drinkers living with HIV/AIDS. *AIDS Behav.* 2008; 12:729–740. [PubMed: 18071894]
9. Abayomi A, Goodridge W, Asika O. Wireless networks for surveillance, data capture and data management in the human immunodeficiency virus epidemic care and treatment programmes. *Afr J Med Med Sci.* 2006; (35 suppl):149–152. [PubMed: 18050790]
10. Barnighausen T, Chaiyachati K, Chimnindi N, Peoples A, Haberer J, Newell ML. Interventions to increase antiretroviral adherence in sub-Saharan Africa: a systematic review of evaluation studies. *Lancet Infect Dis.* 2011; 11:942–951. [PubMed: 22030332]
11. Haberer JE, Kiwanuka J, Nansera D, Wilson IB, Bangsberg DR. Changes in using mobile phones for collection of antiretroviral therapy adherence data in a resource-limited settings. *AIDS Behav.* 2010; 14:1294–1301. [PubMed: 20532605]
12. Sobell, LC.; Sobell, MB. Timeline follow-back: a technique for assessing self-reported ethanol consumption. In: Litten, RZ.; Allen, J., editors. *Measuring Alcohol Consumption: Psychosocial and Biological Methods.* Humana Press; Totowa, NJ: 1992. p. 41-72.
13. Sobell LC, Brown J, Leo GI, Sobell MB. The reliability of the alcohol timeline follow-back when administered by telephone and by computer. *J Drug Alcohol Depend.* 1996; 42:49–54.
14. Hetteema JE, Miller WR, Tonigan JS, Delaney HD. The Form 90-DWI: an instrument for assessing intoxicated driving. *Psychol Addictive Behav.* 2008; 22:117–121.
15. Ingersoll KS, Farrell L, Cohen J, et al. Increasing HAART adherence and reducing cocaine use. A randomized clinical trial of two adherence and drug use interventions for HIV+ cocaine users. *J Drug Alcohol Depend.* 2011; 116:177–187.
16. Arnsten JH, Li X, Mizuno Y, Knowlton AR, Gourevitch MN, Handley K. Factors associated with antiretroviral therapy adherence and medication errors among HIV-infected injection drug users. *J Acquir Immune Defic Syndr.* 2007; 46(suppl 2):S64–71. [PubMed: 18089986]
17. Ingersoll K. The impact of psychiatric symptoms, drug use, and medication regimen on non-adherence to HIV treatment. *AIDS Care.* 2004; 16:199–211. [PubMed: 14676026]
18. Parsons JT, Rosof E, Mustanski B. Medication adherence mediates the relationship between adherence self-efficacy and biological assessments of HIV health among those with alcohol use disorders. *AIDS Behav.* 2008; 12:95–103. [PubMed: 17503172]
19. Lin IKL. A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989; 45:255–268. [PubMed: 2720055]
20. Lin IKL. A note on the concordance correlation coefficient. *Biometrics.* 2000; 56:324–325.
21. Ortego C, Huedo-Medina TB, Llorca J, et al. Adherence to highly active antiretroviral therapy (HAART): a meta analysis. *AIDS Behav.* 2011; 15:1381–1396. [PubMed: 21468660]
22. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979; 86:420–428. [PubMed: 18839484]
23. McBride, GB. NIWA Client Report: HAM2005-062. National Institute of Water & Atmospheric Research; Hamilton, New Zealand: 2005. A proposal for strength-of-agreement criteria for Lin's Concordance Correlation Coefficient.
24. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *Statistician.* 1983; 32:307–317.
25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1:307–310. [PubMed: 2868172]
26. Bland, M. *An Introduction to Medical Statistics.* Oxford University Press; New York: 1995.
27. Hanneman SK. Design, analysis and interpretation of method comparison studies. *Adv Crit Care.* 2008; 19:223–234.
28. Soeken KL, Prescott PA. Issues in the use of Kappa to estimate reliability. *Med Care.* 1986; 24:733–741. [PubMed: 3736144]

29. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
30. Safren SA, Otto MW, Worth JL, et al. Two strategies to increase adherence to HIV antiretroviral medication: life-steps and medication monitoring. *Behav Res Ther*. 2001; 39:1151–1162. [PubMed: 11579986]
31. Arnsten JH, Demas PA, Farzadegan H, Grant RW, Gourevitch MN, Chang CJ. Antiretroviral therapy adherence and viral suppression in HIV-infected drug users: comparison of self-report and electronic monitoring. *Clin Infect Dis*. 2001; 33:1417–1423. [PubMed: 11550118]
32. Nichol MB, Venturini F, Sung JC. A critical evaluation of the methodology of the literature on medication compliance. *Ann Pharmacother*. 1999; 33:531–540.
33. Llabre MM, Weaver KE, Duran RE, Antoni MH, McPherson-Baker S, Schneiderman N. A measurement model of medication adherence to highly active antiretroviral therapy and its relation to viral load in HIV-positive adults. *AIDS Patient Care STDS*. 2006; 20:701–711. [PubMed: 17052140]
34. 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; 10:247–248. [PubMed: 16783536]

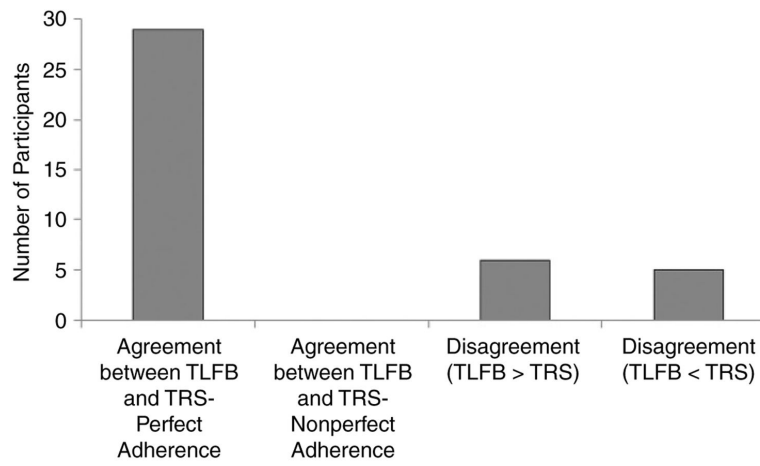


Figure 1. Agreement or disagreement between timeline follow-back (TFLB) and telephone reporting system (TRS) total adherence rates in terms of number of participants.

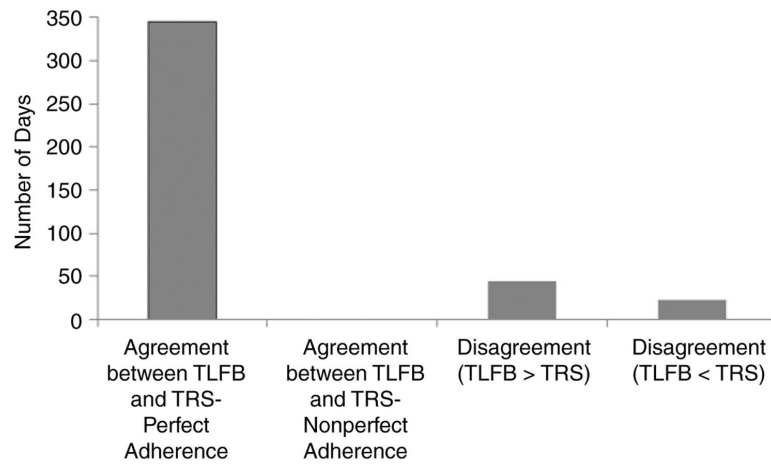


Figure 2. Agreement or disagreement between timeline follow-back (TFLB) and telephone reporting system (TRS) daily adherence rates in terms of number of days.