Published in final edited form as:

Psychol Addict Behav. 2005 September ; 19(3): 291–295. doi:10.1037/0893-164X.19.3.291.

Do Daily Interactive Voice Response Reports of Smoking Behavior Correspond With Retrospective Reports?

Benjamin A. Toll and Ned L. Cooney

Yale University School of Medicine and Veterans Affairs Connecticut Healthcare System

Sherry A. McKee and Stephanie S. O'Malley

Yale University School of Medicine

Abstract

This study focuses on comparing reports of cigarette consumption using an interactive voice response (IVR) system and the timeline follow-back (TLFB) method. As part of a smoking cessation trial, daily reports of cigarette use were collected using an IVR system and TLFB for 7 days after participants attempted to quit smoking. Reports of cigarette consumption were assessed for the entire sample ($N = 381$) and for the subset of nonabstinent participants ($n = 129$ smokers). The results revealed high correlations between reports of smoking behavior, as assessed by IVR and TLFB. Compliance with the TLFB was superior to compliance with the IVR system. This problem should be addressed if researchers wish to use an IVR system in future smoking cessation clinical trials.

Keywords

interactive voice response; timeline follow-back; smoking; tobacco

Long-term abstinence, as measured by point prevalence assessments of smoking behavior, has traditionally been the most common outcome variable used in smoking cessation clinical trials (Fiore et al., 2000; Piasecki & Baker, 2001; Velicer, Norman, Fava, & Prochaska, 1999; Velicer, Prochaska, Rossi, & Snow, 1992). Although point prevalence smoking abstinence is a useful outcome measure, several alternatives have recently been proposed. Reduction in the rate of smoking behavior as an alternative outcome in harm reduction smoking studies (see, e.g., Hughes, 2000, for a review) appears to be a promising approach that will likely receive further study, and in such studies an assessment of daily smoking behavior for the duration of study participation would be preferable to point prevalence measures. The Society for Research on Nicotine and Tobacco (SRNT) formed a work group to review the literature on measures of abstinence in smoking cessation clinical trials (Hughes et al., 2003). In this comprehensive review, the work group's final recommendation was that researchers conducting clinical trials should report multiple measures of abstinence. Clearly, these researchers are advocating that several measures of abstinence be reported in lieu of solely point prevalence, and in these instances (e.g., prolonged abstinence) it would be preferable to measure daily reports of smoking behavior.

Correspondence concerning this article should be addressed to Benjamin A. Toll, Yale University School of Medicine, Department of Psychiatry, 1 Long Wharf Drive, Box 18, New Haven, CT 06511. E-mail: benjamin.toll@yale.edu.
Benjamin A. Toll and Ned L. Cooney, Department of Psychiatry, Yale University School of Medicine and Veterans Affairs Connecticut Health-care System, West Haven, Connecticut; Sherry A. McKee and Stephanie S. O'Malley, Department of Psychiatry, Yale University School of Medicine.

Daily reporting of smoking may provide a more valid assessment of behavior than longer retrospective reports because it involves reporting over a shorter time period. Shiffman et al. (1997) conducted a study focusing on assessments of cigarette consumption and smoking lapse behavior (e.g., mood ratings) in which comparisons were made between real-time reports on hand-held computers and retrospective accounts provided after 12 weeks. In this study, participants' recall of lapses was poor, and they overestimated their negative affect and the number of cigarettes smoked during a lapse. Shiffman et al. (1997) suggested that caution be used with retrospective recall in research.

Interactive voice response (IVR) technology has been shown to reliably assess a variety of daily behaviors (Piette, 2000). With regard to substance use, studies have shown that daily reports of drinking can be adequately assessed using IVR technology (e.g., Searles, Perrine, Mundt, & Helzer, 1995). Although IVR has been used in smoking cessation clinical trials (Ershoff et al., 1999; Killen et al., 2000; Shiffman et al., 2002), the reliability and validity have never been formally evaluated with smokers.

The alcohol timeline follow-back (TLFB) method (Sobell & Sobell, 1992, 1995, 2003) provides a thorough assessment of a client's drinking history for the past 30 to 360 days, and numerous studies have shown satisfactory reliability and validity for this instrument. The TLFB has also been evaluated as a measure of past illicit drug use and tobacco smoking, and it has been shown to have adequate reliability and validity with these substances as well (Brown et al., 1998; Fals-Stewart, O'Farrell, Freitas, McFarlin, & Rutigliano, 2000; Hersh, Mulgrew, Van Kirk, & Kranzler, 1999).

Although there are no studies to date that have compared IVR to TLFB for smoking behavior, several studies have compared assessments of daily drinking as measured with IVR with those measured by TLFB. Bardone, Krahn, Goodman, and Searles (2000) assessed how well these two types of evaluation measured the distribution of the average number of drinks per month for 12 weeks. In a sample of 43 participants, IVR and TLFB identified approximately the same numbers of individuals drinking at set cutoff points (e.g., seven or more drinks/week), showing good convergence between the two measures. Searles, Helzer, and Walter (2000) conducted a study in which IVR and TLFB assessments were compared in a sample of 33 participants. Modest correlations were found between the two measures for quantity of beverages consumed, number of drinking days, and number of heavy drinking days (e.g., aggregate total of $r = .57$ for IVR and TLFB [at 366 days]). In addition, there was broad variability across individuals. In a more fine-grained analysis of this data set, Searles, Helzer, Rose, and Badger (2002) again compared IVR and TLFB measures of alcohol consumption. The TLFB consistently underestimated alcohol consumption compared with IVR (e.g., 3.28 drinks/day reported by IVR compared with 2.75 drink/day reported by TLFB [at 366 days]), and there appeared to be no relationship between the TLFB reports and the length of time for the recall period.

Thus, in the three studies that have compared IVR and TLFB assessments of substance use, measures of association showed considerable variability. In addition to the fact that there are a small number of studies drawing comparisons, all of these studies focused on alcohol use in relatively small samples. In the present study, IVR reports of smoking behavior were compared to TLFB reports of such behavior in a large sample of individuals participating in a smoking cessation clinical trial.

Method

Participants

This report is based on a secondary analysis of data from a double-blind, placebo-controlled clinical trial for smoking cessation that examined whether naltrexone augments the effects of

the nicotine patch (O'Malley, 2004). Participants were eligible for inclusion in the clinical trial if they were at least 18 years of age, smoked 20 cigarettes per day for at least 1 year, and had a baseline expired carbon monoxide (CO) level of at least 10 parts per million (ppm). Participants were excluded if they were medically unstable, suffering from a major psychiatric disorder, or alcohol dependent. This study was approved by the Human Investigation Committee of the Yale University School of Medicine.

As a component of their participation in the clinical trial, all participants agreed to take part in a project involving daily reporting of their cigarette and alcohol use, urge to smoke, nicotine withdrawal, and depressive symptoms. Three hundred eighty-five individuals participated in the clinical trial, but there were 4 study participants who discontinued their study participation after their first treatment session and failed to provide any subsequent data with regard to their tobacco use. The present analyses are based on the 381 individuals who participated in the study treatment and provided one or more data points via the IVR system or by TLFB. The analyses in this article also focus on study participants who smoked ($n = 129$) during their study participation. Demographic characteristics for the entire sample and the smoker subset are presented in Table 1.

Procedure

As part of a smoking cessation clinical trial, participants received a 21-mg nicotine patch daily and were randomized to placebo or to receive 25, 50, or 100 mg of naltrexone for a 6-week period. For 6 weeks after their quit date, participants came to the clinic for research appointments on a weekly basis. Questionnaires were administered, and brief counseling was provided at each of these appointments. Smoking was assessed using TLFB methodology at each weekly appointment using procedures outlined by Sobell and Sobell (1992, 1995, 2003). Specifically, baseline TLFB data were gathered for the 30 days prior to their first screening session and at all weekly appointments individuals were verbally asked to indicate the number of cigarettes they consumed each day for the preceding week. The TLFB has been shown to be a reliable and valid method of assessment of smoking behavior (Brown et al., 1998), and it has been shown to have good convergence with the biochemical markers CO and cotinine (Brandon, Copeland, & Saper, 1995).

Smoking status was biochemically verified by measuring CO using a Vitalograph BreathCO monitor (Lenexa, KS), with readings of 10 ppm or greater coded as current smoking (SRNT Subcommittee on Biochemical Verification, 2002). This biomarker is adequate for the assessment of smoking as "any specimen in which the biomarker can be measured is suitable for determining smoking status" (SRNT Subcommittee on Biochemical Verification, 2002, p. 150). In addition, the sensitivity and specificity of this biomarker are both approximately 90%, which is slightly lower than cotinine (96%–97% sensitivity and 99%–100% specificity).

The present study is based on data gathered pertaining to the 1st week after participants' quit date. Participants were instructed to call the IVR telephone line first thing in the morning for 7 days in a row, beginning the day after their quit date. Participants called a toll-free number that was active throughout the United States, which was connected to a computerized system that collected daily reports for each of the 7 days. Participants responded to a brief series of questions¹ about their smoking, alcohol consumption, mood, and nicotine withdrawal symptoms by touching numbers on their telephone keypad that corresponded to their answers. Data were consequently entered directly into a database for later evaluation and analysis. The first 11 questions were preceded by a keyword, following which answers to questions could be immediately entered. This eliminated the need for participants to listen to each question in

¹The full IVR questionnaire is available upon request from Benjamin A. Toll.

its entirety after they had experience using the system. To increase the compliance with daily dairies, for each day the telephone questionnaire was completed, participants earned \$2. Participants were not paid specifically to provide the TLFB and questionnaire data included in the present analyses; however, they were paid \$100 if they completed the treatment and follow-up appointments for the smoking cessation clinical trial.

Results

IVR Response Rate and TLFB Data Collection for the Entire Sample

The 381 participants were asked to report their cigarette and alcohol consumption using the IVR system on 7 days, for a total of 2,667 possible reports. In fact, 84.0% of these requested reports were made. Complete data (i.e., reporting every day over the entire 7-day interval) were obtained for 49.9% of the sample. It should be noted that the number of daily reports either stayed the same or progressively declined each successive day, with a precipitous decline on the 7th day. The percentages of complete data (i.e., reporting every day) for each of the 7 days, beginning with Day 1, were 94.2%, 86.4%, 80.6%, 74.8%, 70.1%, 65.1%, and 49.9%. The larger drop-off on the 7th day may have been due to the structure of the clinical trial. Specifically, participants were seen for their second treatment/research session on their 6th day of calling the IVR system, and even though they were told to call on the day after this session, it seems that many of them forgot or disregarded this instruction, perhaps assuming they had already reported on their smoking the day before.

As with the IVR system, there was an attempt to collect 2,667 possible reports (i.e., 381 participants reporting over 7 days) using TLFB. From these, a total of 96.3% of reports were obtained. Complete data (i.e., reporting over the entire 7-day interval) were obtained for 93.2% of the sample.

Sensitivity and Specificity for the Entire Sample

At the time of the second treatment session of the clinical trial, participants' expired air CO was measured. This measurement was considered to be the gold standard and was compared with participants' IVR and TLFB reports of cigarette use for that day to calculate sensitivity and specificity indices. A CO value of 10 ppm or greater was classified as indicative that the participant had been smoking. The *sensitivity* of a test is the proportion of truly disordered individuals in a screened population who are classified as disordered by the screening test (i.e., true positives). For example, with smoking behavior, a test would have good sensitivity if it correctly identified individuals as smoking when they were truly smoking, as confirmed by their CO measurement. The *specificity* of a test is the proportion of truly nondisordered people in a screened population who are identified as such by the screening test (i.e., true negatives). For example, with smoking, a test would have good specificity if it correctly identified individuals as not smoking when they were truly not smoking. *Positive predictive value* refers to the proportion of test-positive individuals who have the disorder, and *negative predictive value* refers to the proportion of test-negative individuals who do not have the disorder.

Both the IVR and TLFB appeared to have poor sensitivity and positive predictive value, with both values equaling 0 for IVR and values of .3 and .18, respectively, for TLFB. In contrast, both the IVR and TLFB had good specificity and negative predictive value, with values of .91 and .98, respectively, for IVR and values of .93 and .97, respectively, for TLFB. In other words, both measures failed to identify smokers but accurately identified most nonsmokers. The concordance between the IVR and CO measures and between the TLFB and CO measures are presented in Tables 2 and 3, respectively.

As can be seen in Tables 2 and 3, it is important to note that the indices of sensitivity and positive predictive value may be spuriously low because of the small percentage of participants included in these analyses (i.e., very few participants reported smoking). Put another way, these results were likely affected by the fact that these individuals were assessed shortly after their target quit date in a smoking cessation clinical trial such that most participants reported nonsmoking, with a much smaller proportion reporting smoking. This finding also is a function of the CO cutpoint that was used in this study. Last, there may be measurement error in this analysis because of the fact that CO was used as the method of biochemical validation, and CO has a half-life of 2 hr (SRNT Subcommittee on Biochemical Verification, 2002). It would have been more accurate to report cotinine levels as the method of biochemical validation because cotinine has a half-life of 16 hr, but this biochemical measure was not collected.

Smoking

Correlations between IVR and TLFB reports of smoking for each of the 7 assessment days and an aggregate total (i.e., the sum of the values reported for Days 1–7) are presented in Table 4. As reported in Table 4, all correlations are significant and fall within the high range. The sample size decreased nearly every reporting day, and on most days the correlation decreased as well.

We analyzed concordance between TLFB and IVR measurement methods for the subsample of 129 participants who reported smoking on at least 1 day during the 7-day time interval. Of the 903 possible reports that those participants were requested to make by both TLFB and IVR, 670 corresponding data pairs were available. Of those 670, there were 54 instances in which the TLFB value was equal to zero and the IVR value was greater than zero, and the range for these values was from 1 to 20 cigarettes, with a mean of 3.28 ($SD = 3.94$). There were 36 instances in which the IVR value was equal to zero and the TLFB value was greater than zero, and the range of these values was from 0.5 to 25 cigarettes, with a mean of 2.47 ($SD = 4.22$). Concordance for reports of abstinence or nonabstinence from cigarettes between the TLFB and IVR cigarette consumption values for the 670 recorded data pairs is presented in Table 5.

Discussion

Comparisons were made between daily reports of cigarette consumption as measured by an IVR system and the TLFB. Although such comparisons have been made previously for alcohol use, this is the first study to do so for cigarette smoking. One of the most important findings of this study is that less complete data were obtained using the IVR system compared with the TLFB. Although 84% of required IVR assessments were obtained, complete data (i.e., reporting every day over the 7-day time interval) were obtained for 49.9% of the sample for IVR compared with 93.2% of the sample for TLFB. In addition, the sample size decreased for each day that behavior was reported. This may represent a problem for researchers who wish to conduct smoking research in the future using an IVR system.

These findings of low compliance with IVR reporting are similar to those reported in a recent large-scale smoking cessation clinical trial (Shiffman et al., 2002) in which less than 10% of participants provided IVR reports every day for the first 3 weeks of the clinical trial. In the present study, research assistants collected smoking data from participants when the participant attended a treatment session. This is therefore a potential strength of the TLFB, as it ensures data collection for each participant who does not drop out of study treatment. However, both the current study and Shiffman et al.'s (2002) study were conducted during the treatment phase of a clinical trial. Future studies should assess the utility of using an IVR system during the follow-up phase of a clinical trial, when it is more difficult to get participants to attend research sessions. Participants may be more willing to call a toll-free number to access an IVR system than to attend a research interview. In particular, participants who have dropped out of a clinical

trial may be more willing to provide follow-up data if offered the option to call an IVR system and given incentives to do so (e.g., a monetary payment).

It should also be noted that previous studies that have found much better compliance with IVR systems for the reporting of drinking behavior used greater monetary incentives and reminder telephone calls (Bardone et al., 2000; Searles et al., 2000). Indeed, an IVR study conducted by Searles et al. (2000) included a bonus of \$500 if participants completed 95% of their calls on time and attended all quarterly interviews, and this study also incorporated a competition for a total of \$6,000 to be divided among the three best IVR records. If one takes this into account, compliance with the IVR system regarding both the frequency of complete data and the number of participants reporting per day in the present study may have been better if greater incentives and reminder phone calls were used. Conversely, because the response rate was already low, the number of reports may have been even less if monetary incentives had not been used at all. This concern was raised by Sobell and Sobell (2003), who noted that the level of compliance that would be found without monetary incentives is uncertain. This is an important point of caution for researchers who wish to use an IVR system in the future.

Sobell and Sobell (2003) discussed other concerns about using an IVR system to assess drinking behavior. Even though the present article focuses on smoking behavior, several of these concerns are pertinent to the current discussion. Sobell and Sobell (2003) stated that none of the studies comparing IVR with TLFB have used a clinical population. In addition, there have been no reports of IVR being superior to self-report, which is much less costly. Finally, the only method to assess pretreatment data prospectively would be for participants to self-monitor before they began treatment. This procedure has two major problems. First, it would mean that treatment would need to be delayed solely to gather pretreatment data prospectively. Second, the very act of self-monitoring might be reactive, which may call into question whether the data are representative of pretreatment behavior (Sobell & Sobell, 2003).

In the current study, both IVR and TLFB were found to have poor sensitivity and positive predictive value and good specificity and negative predictive value. They both failed to identify smokers but accurately determined to a high degree that participants were not smoking, as verified by the CO measure. Because these data were collected shortly after the target quit date in a smoking cessation trial, most participants reported not smoking, which may account for these results. However, this represents a potentially serious problem, as indices of sensitivity and positive predictive value should be in the high range even when base rates of the disorder are low. Future smoking cessation clinical trials should investigate this subject using a biochemical marker that has a longer half-life than CO (e.g., cotinine).

With regard to the results from the smoking subsample, there were high correlations between IVR and TLFB for both day-by-day and aggregate total measurement, but the number of reports and the correlation decreased nearly every day. Reporting biases in cigarette consumption (i.e., reporting no cigarette use on IVR and reporting smoking on the TLFB, or vice versa) were approximately equal for both types of measurement. Although the range of values was greater on the TLFB, there was a slightly higher percentage of the number of reports of smoking days and a higher mean number of cigarettes reported on the IVR.

There are several limitations to this study. First, participants may have made more accurate reports on the TLFB because they had used the IVR reporting system. Second, with larger monetary incentives and reminder calls, compliance with the IVR system might have been better. Third, compliance may have also been different if participants were required to report to the IVR system throughout the study treatment instead of only 1 week. Fourth, the TLFB procedure asked participants to recall 1 week of smoking. Longer periods of TLFB retrospection may yield different results. Fifth, this report has focused only on measures of

smoking behavior. Other processes, such as mood states or withdrawal symptoms, may be measured more accurately using the smaller retrospective timeframe of the IVR methods. Finally, caution is warranted regarding the generalization of these findings given that the participants in this study were unusual in several respects. Specifically, these participants were paid for completing IVR ratings as part of a clinical trial involving medications. Also, participants came to a clinic weekly, so reports were easily gathered on a weekly basis rather than at follow-up.

Both IVR and TLFB appear to adequately measure smoking behavior. Although the reporting biases were about equal for each measure, far less complete data were obtained using the IVR system. This may be a critical problem with IVR assessment in gathering data during the treatment phase of a smoking cessation clinical trial. Several other issues should be noted. The ease of administration of the TLFB was superior to IVR; however, the fact that the data were entered directly into a secure database made data management much easier with the IVR system. Even though the IVR system was expensive to administer, new technologies could make IVR assessment more cost effective. Although there are positive and negative qualities for each measure, at present, it appears that TLFB may be a more useful assessment measure of cigarette use for the majority of smoking cessation clinical trials.

Acknowledgements

This research was supported in part by the Office of Academic Affiliations, Veterans Affairs Special Mental Illness Research, Education, and Clinical Centers Fellowship Program in Advanced Psychiatry and Psychology, Department of Veteran Affairs, and by National Institutes of Health Grants K02-AA00171, P50-DA13334, and R01-AA11197. Portions of this article were presented at the 2003 annual convention of the Research Society on Alcoholism in Fort Lauderdale, FL.

References

- Bardone AM, Krahn DD, Goodman BM, Searles JS. Using interactive voice response technology and timeline follow-back methodology in studying binge eating and drinking behavior: Different answers to different forms of the same question? *Addictive Behaviors* 2000;25:1–11. [PubMed: 10708315]
- Brandon TH, Copeland AL, Saper ZL. Programmed therapeutic messages as a smoking treatment adjunct: Reducing the impact of negative affect. *Health Psychology* 1995;14:41–47. [PubMed: 7737072]
- Brown RA, Burgess SA, Sales SD, Whiteley JA, Evans DM, Miller IW. Reliability and validity of a smoking timeline follow-back interview. *Psychology of Addictive Behaviors* 1998;12:101–112.
- Ershoff DH, Quinn VP, Boyd NR, Stern J, Gregory M, Wirtschafter D. The Kaiser Permanente prenatal smoking-cessation trial: When more isn't better, what is enough? *American Journal of Preventive Medicine* 1999;17:161–168. [PubMed: 10987630]
- Fals-Stewart W, O'Farrell TJ, Freitas TT, McFarlin SK, Rutigliano P. The timeline followback reports of psychoactive substance use by drug-abusing patients: Psychometric properties. *Journal of Consulting and Clinical Psychology* 2000;68:134–144. [PubMed: 10710848]
- Fiore, MC.; Bailey, WC.; Cohen, SJ.; Dorfman, SF.; Goldstein, MG.; Gritz, ER., et al. Treating tobacco use and dependence: Clinical practice guideline. U.S. Department of Health and Human Services; Rockville, MD: 2000.
- Hersh D, Mulgrew CL, Van Kirk J, Kranzler HR. The validity of self-reported cocaine use in two groups of cocaine abusers. *Journal of Consulting and Clinical Psychology* 1999;67:37–42. [PubMed: 10028207]
- Hughes JR. Reduced smoking: An introduction and review of the evidence. *Addiction* 2000;95:S3–S7. [PubMed: 10723815]
- Hughes JR, Keely JP, Niaura RS, Ossip-Klein DJ, Richmond RL, Swan GE. Measures of abstinence in clinical trials: Issues and recommendations. *Nicotine & Tobacco Research* 2003;5:13–25. [PubMed: 12745503]

- Killen JD, Fortmann SP, Schatzberg AF, Hayward C, Sussman L, Rothman M, et al. Nicotine patch and paroxetine for smoking cessation. *Journal of Consulting and Clinical Psychology* 2000;68:883–889. [PubMed: 11068974]
- O'Malley, SS. A dose ranging study of naltrexone augmentation of transdermal nicotine patch: Effects on smoking, weight gain, and heavy drinking.. Paper presented at the annual meeting of the Society for Research on Nicotine and Tobacco; Scottsdale, AZ. Feb. 2004
- Piasecki TM, Baker TB. Any further progress in smoking cessation treatment? *Nicotine & Tobacco Research* 2001;3:311–323. [PubMed: 11694198]
- Piette JD. Interactive voice response systems in the diagnosis and management of chronic disease. *American Journal of Managed Care* 2000;6:817–827. [PubMed: 11067378]
- Searles JS, Helzer JE, Rose GL, Badger GJ. Concurrent and retrospective reports of alcohol consumption across 30, 90 and 366 days: Interactive voice response compared with the timeline follow back. *Journal of Studies on Alcohol* 2002;63:352–362. [PubMed: 12086136]
- Searles JS, Helzer JE, Walter DE. Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychology of Addictive Behaviors* 2000;14:277–286. [PubMed: 10998953]
- Searles JS, Perrine MW, Mundt JC, Helzer JE. Self-report of drinking using touch-tone telephone: Extending the limits of reliable daily contact. *Journal of Studies on Alcohol* 1995;56:375–382. [PubMed: 7674671]
- Shiffman S, Dresler CM, Hajek P, Gilbert SJA, Targett DA, Strahs KR. Efficacy of a nicotine lozenge for smoking cessation. *Archives of Internal Medicine* 2002;162:1267–1276. [PubMed: 12038945]
- Shiffman S, Hufford M, Hickcox M, Paty JA, Gnys M, Kassel JD. Remember that? A comparison of real-time versus retrospective recall of smoking lapses. *Journal of Consulting and Clinical Psychology* 1997;65:292–300. [PubMed: 9086693]
- Sobell, LC.; Sobell, MB. Timeline follow-back: A technique for assessing self-reported alcohol consumption.. In: Litten, R.; Allen, J., editors. *Measuring alcohol consumption*. Humana Press; Rockville, MD: 1992. p. 207-224.
- Sobell, LC.; Sobell, MB. Alcohol consumption measures.. In: Allen, JP.; Columbus, M., editors. *Assessing alcohol problems: A guide for clinicians and researchers*. National Institute on Alcohol Abuse and Alcoholism; Bethesda, MD: 1995. p. 55-73.
- Sobell, LC.; Sobell, MB. Alcohol consumption measures.. In: Allen, JP.; Wilson, VB., editors. *Assessing alcohol problems: A guide for clinicians and researchers*. 2nd ed.. National Institute on Alcohol Abuse and Alcoholism; Bethesda, MD: 2003. p. 75-99.
- Society for Research on Nicotine and Tobacco Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. *Nicotine & Tobacco Research* 2002;4:149–159. [PubMed: 12028847]
- Velicer WF, Norman GJ, Fava JL, Prochaska JO. Testing 40 predictions from the transtheoretical model. *Addictive Behaviors* 1999;24:455–469. [PubMed: 10466842]
- Velicer WF, Prochaska JO, Rossi JS, Snow MG. Assessing outcome in smoking cessation studies. *Psychological Bulletin* 1992;24:23–41. [PubMed: 1539088]

Table 1

Sample Demographics

Characteristic	Total (<i>N</i> = 381)	Smokers (<i>n</i> = 129)
Age (years; <i>M</i> [<i>SD</i>])	45.98 (11.15)	45.20 (10.62)
Sex (% male)	51.7	48.1
Marital status (% married) ^a	54.4	42.5
Race (% Caucasian) ^b	87.9	86.1
Employed (% full time) ^c	73.6	75.4

Note. All differences are nonsignificant by *t* test for continuous variables and chi-square for categorical variables.

^a Six cases failed to report any data for this variable.

^b Twenty-four-cases failed to report any data for this variable.

^c Twenty-five cases failed to report any data for this variable.

Table 2

Concordance Between the Interactive Voice Response (IVR) and Carbon Monoxide (CO) Measures

CO test	IVR	
	Smoking	Abstinent
Smoking (%)	0.0	9.0
Abstinent (%)	1.9	89.0

Note. $N = 154$. Values do not add up to exactly 100% because of rounding.

Table 3

Concordance Between the Timeline Follow-Back (TLFB) and Carbon Monoxide (CO) Measures

CO test	TLFB	
	Smoking	Abstinent
Smoking (%)	1.4	6.3
Abstinent (%)	3.2	89.2

Note. $N = 222$. Values do not add up to exactly 100% because of rounding.

Table 4

Correlations Between Interactive Voice Response and Timeline Follow-Back Data for Each Assessment Day and an Aggregate Total for Smoking Behavior

Day	<i>n</i>	<i>r</i>
Day 1	111	.84
Day 2	104	.95
Day 3	100	.93
Day 4	99	.93
Day 5	99	.92
Day 6	87	.88
Day 7	70	.87
Total	46	.99

Note. All correlations are significant ($p < .05$).

Table 5

Concordance Between the Interactive Voice Response (IVR) and Timeline Follow-Back (TLFB) Cigarette Abstinence Reports

IVR	TLFB	
	Smoking	Abstinent
Smoking (%)	39.0	8.1
Abstinent (%)	5.4	47.6

Note. $N = 129$. Table values are based on 670 data pairs. Values do not add up to exactly 100% because of rounding.