How many cigarettes did you smoke? Assessing cigarette consumption by global report, time-line follow-back, and ecological momentary assessment

Saul Shiffman, Ph.D.
Research Department of Psychology, Psychiatry, and Pharmaceutical Science, Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania

Abstract

Objective—This study evaluated and compared several methods of assessing daily cigarette consumption.

Design—Comparison of measures of daily cigarette consumption from several sources, from 232 smokers entering a smoking cessation program.

Main Outcome Measures—Global reports of average smoking, Time-Line Follow-Back recall for the week preceding the study (pre-monitoring TLFB), two weeks' cigarette recordings using electronic diaries and ecological momentary assessment (EMA), and TLFB recall of smoking during EMA (monitored TLFB).

Results—Global reports and pre-monitoring TLFB showed severe digit bias: 6 times as many values as expected were rounded at 10. Monitored TLFB also showed substantial digit bias (4 times). EMA data showed none. EMA averaged 2.6 cigarettes lower than monitored TLFB, but exceeded TLFB on 32% of days. Across days, EMA and TLFB only correlated 0.29. Daily variations in TLFB did not correlate with variations in carbon monoxide (CO) measures taken on three days, but EMA measures did; among subjects whose CO varied, r= 0.69. CO correlated with EMA cigarettes recorded in the preceding two hours, suggesting timely recording of cigarettes.

Conclusion—TLFB measures are limited for precise assessment of cigarette consumption. EMA measures appear to be useful for tracking smoking, and likely other health-relevant events.

Keywords

smoking; digit bias; ecological momentary assessment

In studying smokers and smoking, one of the most fundamental assessments is determining how much a person smokes. This is almost always reported as a descriptor of samples in smoking studies, and is very often analyzed as a predictor in studies of tobacco dependence, smoking patterns, and cessation outcome. In almost all cases, daily cigarette consumption is assessed by global self-report; that is, by asking subjects how much they smoke, on average.

Global assessments of smoking rates might be challenged on several grounds. Studies of cognitive processes that people use to answer survey questions show that respondents do not engage in systematic counting or other precise processes, but rather base their answers on broad cognitive heuristics, which can introduce both inaccuracy and bias (Hammersley, 1994; Bradburn, Rips, & Shevell, 1987; Shiffman et al., 2008). One common indicator of the relatively crude process used to generate such estimates is "digit bias" or "heaping" – the tendency of the estimates to cluster around rounded value (e.g., Nagi, Stockwell, & Snavley, 1973; Shryock & Siegel, 1976). This has been observed for reports of many health-related items, such as current age, age of menopause, blood pressure, height and weight, and other quantities, and is seen as a significant limitation of such self-reports (Hessel, 1986; Rowland, 1990; Crawford, Johannes, & Stellato, 2002, UN, 2003). Digit bias has been observed in reported smoking rates in a national sample of adult smokers (Klesges, Debon, & Ray, 1995) and also among adolescents (Lewis-Esquerre et al., 2005). Because cigarettes are packaged in packs of 20, one might think that some heaping is due to true clustering of daily consumption around that number, or multiples thereof. However, Klesges et al. (1995) examined a biochemical indicator of smoking and found that, in contrast to the self-report data, it showed no tendency towards "heaping" at particular values, suggesting that the heaping observed in self-report was due to digit bias and not to an actual tendency for daily consumption to cluster at rounded values representing packs.

Global assessments of daily consumption also gloss over variability in cigarette consumption over time, which might be particularly important for understanding smoking patterns (Chandra et al., 2007) and for studying systematic changes in consumption (Hatsukami et al., 2007), which have become increasingly important with increased emphasis on smoking reduction as a means of reducing the harm of smoking or as an approach to quitting (Hughes & Carpenter, 2006). One approach that has been suggested for collecting more refined and more accurate self-reports of smoking is the Time-Line Follow-Back (TLFB), which asks subjects to retrospectively report daily cigarette consumption day-by-day over some period of time. The TLFB has been used widely to assess alcohol consumption, with data often collected based on recall over months (Monti et al., 2007; Miller & Del Boca, 1994; Epstein et al., 2004). One study (Lewis-Esquerre et al., 2005) evaluated the TLFB method with adolescent smokers, and found it produced less digit bias than global reports, and that averaged consumption figures from TLFB correlated reasonably well with biochemical measures of smoking. The one study that evaluated TLFB in adult smoking found good cross-subject agreement between 1-week TLFB and end-of-day reports of cigarette consumption (Toll, Cooney, McKee, & O'Malley, 2005). That study but did not evaluate the correspondence between the methods in assessing smoking across days and did not assess heaping, which could have pervaded both measures, since end-of-day recall is also subject to some recall bias.

Beyond just tracking the amount someone smokes on average or on a particular day, self-monitoring methods are important for documenting patterns of smoking within a day (Chandra et al., 2007) and the relationship of smoking to situational antecedents. Such data have be obtained by having subjects record cigarettes in real time, using the methods of EMA (Stone & Shiffman, 1994; Shiffman et al., 2008), which calls for real-time data collection in real-world settings. Both Colby and Klesges suggest that Global and TLFB data be compared against EMA data. EMA methods have been used to document within-day patterns of smoking (Chandra et al., 2007), and to compare the situations associated with smoking with those in which smoking does not occur (Shiffman et al., 2002). These analyses (Shiffman et al 2002; Shiffman et al., 2004) have demonstrated expected associations, for example, between alcohol and smoking and between craving and smoking but have also yielded some surprising findings, notably the lack of an association between smoking and mood during ad lib smoking.

Importantly, the ability to estimate such associations depends in part on smokers' compliance with the instruction to faithfully track their smoking. The smoker's task in these studies is seemingly daunting, as they are asked to make dozens of entries per day as they go about their daily life. While EMA studies using appropriately-designed electronic diaries have demonstrated excellent compliance with response to audible prompts issued by the diary – often achieving better than 90% compliance (Shiffman et al., 2002; Hufford & Shields, 2002) – it is more difficult to evaluate compliance with the monitoring of cigarettes or other behaviors, which must occur at the subjects' initiative.

The purpose of this paper is to examine and compare self-reports of cigarette consumption of three types: global reports, TLFB and EMA. Within TLFB reports, we examine TLFB reports from two periods: reports made while smokers were tracking their smoking using EMA, and reports collected in the more typical way, when smokers were not otherwise tracking their smoking. We assess digit bias in each method, and compare and correlate the estimates of cigarette consumption from each, both across subjects and across days within subjects, to examine the day-by-day association between TLFB and EMA.

Beyond comparing the different modes of self-report, we also attempted to validate each via their association with two objective biochemical indices of smoke exposure (Benowitz et al., 2002). Cotinine is the major metabolite of nicotine, which is rather specifically absorbed from smoking. Cotinine's half-life is 16 hours, so it reflects nicotine absorption over the preceding few days (Benowitz et al., 2002). We examined its association with reported cigarette consumption (EMA and TLFB) in the three days preceding the cotinine measure. Carbon monoxide (CO) is a by-product of incomplete combustion, is absorbed during smoking, and can be assessed in exhaled breath. We used multiple assessments of CO to assess the sensitivity of TLFB and EMA data to capture between-day variations in smoking. CO has a half-life of 2 hours (Benowitz et al., 2002), and thus is particularly sensitive to very recent smoking. Accordingly, CO levels provided a way to test the timeliness of smoker's recording of cigarettes in EMA, by correlating CO with the number of cigarettes recorded in the preceding two hours. The relationship between smoking and both CO and cotinine is modest (e.g., Swan et al., 1993), because the amount that is absorbed from smoking and the rate at which clearance occurs (as well as input from other sources [for CO] and rate of production [cotinine], Benowitz et al, 2002) are subject to variability. Nevertheless, these measures are biologically linked to smoking and have demonstrated relationships to consumption, and thus provide objective indices of smoke exposure against which self-reports can be compared.

Methods

Subjects

Participants were 232 smokers who enrolled in a smoking cessation research study. To qualify for the present analyses, participants had to have data from at least 7 days on both EMA and TLFB during Days 2–15 of monitoring. Participants were recruited through advertisements for smoking cessation treatment and paid $50 for participating. To qualify, participants had to report smoking at least 10 cigarettes per day, to have been smoking for at least 2 years, and to report high motivation and overall efficacy to quit during a screening interview (combined score of 150 on the sum of two 0–100 scales). Data from this study have been used in other publications (e.g., Shiffman et al., 2002).

The mean age of the sample was 43.6 (9.8) years. The sample was majority female (59%) and Caucasian (92%). On questionnaires, participants reported smoking for 22.9 (10) years, and smoking their first cigarette of the day 16.5 (27.1) minutes after waking.

Procedures

After informed consent was obtained, participants completed a smoking history, including an item that asked how many cigarettes they smoked each day, on average (Global reports). At this time, participants also completed a TLFB assessment of the preceding week, which we designate as the pre-monitoring TLFB. (Collection of this TLFB was initiated after the study was under way, so was only available from 194 participants who had pre-monitoring TLFB data for at least 4 days.) Participants were then trained to use a hand-held computer designed to allow data collection in near real time. The electronic diary (ED) was programmed on a PSION Organiser II LZ 64 (PSION, Ltd., London, England), a handheld computer with a 4-line, 20-character LCD screen, $5.6 \times 3.1 \times 1.1$ in., weighing 8.8 oz. Using ED, participants monitored ad-lib smoking for 16 days prior to a designated quit date; they were instructed not to change their smoking during this time. (Previous analyses of CO levels (Shiffman et al, 2002) showed no significant change over this period.) The first day of monitoring was incomplete and therefore excluded. We used the following two weeks (days 2–15), which ended two days before the target quit date.

During the monitoring period, participants were instructed to record each cigarette on the ED, immediately before smoking. On most of these occasions, participants had only to push a single key; ED simply recorded the smoking event. On about 4–5 smoking occasions per day, selected at random by ED, ED administered an assessment. In addition to recording cigarettes in real time, participants had an opportunity each evening to report whether they had smoked any cigarettes that they failed to record, and how many went unrecorded. Cigarettes reported in this way amounted to 3.8% (SD=5.3%) of subjects' total daily entries. When subjects arrived for a clinic visit, their ED's were taken from them for data transfer and hardware checks. At the end of the visit, subjects recorded any cigarettes they had smoked during the visit. These cigarettes were counted towards their daily EMA total (EMA CPD). In addition to recording cigarettes, participants were also prompted audibly by the ED 4–5 times daily at random times to complete an assessment while they were not smoking. Participants responded to 91% of all prompts within the 2 min allowed, indicating that they were carrying and attending to ED. Monitoring and prompting covered all waking hours (the ED included an alarm clock function so that participants could put ED to sleep and avoid prompting while they slept). Subjects had the option to suspend prompting for a limited time when they could not be disturbed; they could record cigarettes during these times and when driving, without being subject to assessment, but were not required to.

At Days 8 and 15, subjects returned to the clinic and completed a TLFB assessment indicating how many cigarettes they had smoked on each day going back to the previous visit, typically 5 or 7 days prior. The TLFB form was in the form of a calendar, and subjects were encouraged to use personal cues and events as aids to recall. We refer to this TLFB obtained during EMA monitoring as "monitored TLFB" or simply "TLFB."

Participants received nonpharmacological psychoeducational treatment in groups of 8 to 16 at clinic visits typically falling on Days 1, 3, 8, and 15. Sessions were generally held in the evening at about 6 PM; three groups (22 participants) had sessions in the morning (10 am or noon). (Smokers assessed in the morning demonstrated CO levels 6 ppm lower, but including time of assessment in analyses did not affect the results at all, so we present the simpler models based on combined data.)

Biochemical measures

At clinic visits 1 – 4 during the baseline period, Days 1, 3, 8, and 15, subjects provided a breath sample for CO assessment (in parts per million – ppm); we used data from days 3, 8, and 15 for analysis, as these days overlapped with the EMA and monitored TLFB data being

analyzed. (195 subjects had CO values and smoking data for all 3 measurement occasions, and an additional 35 subjects had data for 2 (n=28) or 1 (n=7) days). On Day 8 (Visit 3), participants also provided a saliva sample that was subsequently assayed for cotinine concentration (in ng/ml); cotinine values were available for 208 subjects. (The days of sample collection varied somewhat with clinic schedules over the course of the study, and in individual instances where participants could not attend a scheduled visit but were rescheduled for a different day.)

Analyses

We analyzed the TLFB and EMA data at both the subject and day level. We excluded 2 participants, who initially reported smoking 80 and 90 cigarettes per day, respectively, and who were consistent outliers in the analyses (the results were largely unaffected).

To assess digit bias, we noted whether the values reported for Global CPD, daily TLFB, and daily EMA were even multiples of 10. (We also tracked even multiples of 5, but the results were conceptually similar, so we concentrate on heaping at intervals of 10.) For each participant, we computed the proportion of days demonstrating digit bias, and expressed this using Whipple's index (Denic et al., 2004), which indicates the factor by which the proportion of "rounded" values exceeds that expected from a random distribution of values.

We correlated EMA and TLFB measures with CO and cotinine assays obtained during the monitoring period. We tested for and saw no consistent evidence of curvilinear relationships (cf. Swan et al., 1993), so we focus on linear associations. In addition to univariate models, we also tested which measure was associated with unique variance in multivariate models that included both EMA and TLFB. To assess the measures' sensitivities to variations in smoking, we used hierarchical regression methods (with subjects as random effects) to assess the within-subjects relationship between daily smoking and CO across 3 CO assessments. Finally, to assess the timeliness of cigarette entries within a day on EMA, we focused the analysis on cigarette records in the two hours preceding the CO assessments, correlating the number of cigarettes entered in that time with the observed CO level.

Results

Over an average of 13.7 (0.9) days, each subject recorded an average total of 303.5 (SD=122.5) individual cigarettes on ED, with an additional 9.2 (11.1) reported at end of day. Each subject completed 59.7 (18.2) smoking assessments, and an additional 62.6 (20.2) randomly-scheduled assessments when they were not smoking. A total of 341.4 (137.3) cigarettes were reported by TLFB.

Digit Bias

Figure 1 shows the distribution of values of CPD for Global reports, pre-monitoring and monitored TLFB, and EMA reports. Digit bias is evident in the Global and TLFB distributions. As shown in table 1, Global reports and pre-monitoring TLFB reports were round multiples of 10 more than 6 times as often than expected, and daily monitored TLFB values were rounded to 10 over 4 times more often than expected (all p<0.0001). In contrast, daily EMA values fit the expected random distribution almost exactly. (The distributions also show peaks at even multiples of 5, capturing 85%, 87%, 66%, and 20% of daily Global, pre-monitoring TLFB, monitored TLFB, and EMA values, respectively.)

As shown in Table 1, the pre-monitoring TLFB data were also marked by invariant reporting. A majority of subjects (51.7%) reported exactly the same cigarette consumption for each of the days preceding the TLFB report, typically entering uniformly the number of cigarettes they had indicated in their global report (81% of those reporting unvarying rates).

In contrast, only 3% of monitored TLFB records and 0% of EMA indicated uniform cigarette consumption each day.

Subjects who reported rounded values for Global CPD were also more likely to report rounded values on the pre-monitoring TLFB (point-biserial r=0.63, p<.0001), and the monitored TLFB (point-biserial r=0.37, p<.0001), but EMA values were unrelated to rounding of Global or TLFB values (all r<0.09, ns; results were similar for multiples of 5).

Reported daily cigarette consumption

Table 1 shows the average daily smoking rate assessed by Global, TLFB (pre-monitoring and monitored), and EMA. Global reports of cigarette consumption were almost identical to those obtained by pre-monitoring TLFB. However, monitored TLFB reports averaged about 1.5 cigarettes per day lower than Global reports, and EMA reports averaged about 2.5 cigarettes lower than TLFB reports for the same period. Each of these differences was statistically significant. For the period where TLFB and EMA overlapped, subjects' daily EMA cigarette consumption averaged 92.6% (SD=18.8%) of their daily TLFB cigarette consumption.

Although the aggregated EMA smoking rate was, on average, lower than that reported on monitored TLFB, this was not consistently the case for all subjects, nor across days. For 30% of all subjects, the average CPD by EMA was higher than that calculated by contemporaneous TLFB. Moreover, the vast majority of subjects (87%) had at least some days when they recorded more cigarettes on EMA that they later recalled on TLFB. Across subjects, this was true on 31.5% (SD=26.1%) of days, suggesting that the lower average EMA smoking rate cannot be solely due to failure to record cigarettes on EMA.

As seen in Figure 1, the most common daily cigarette consumption reported by monitored TLFB was 20 cigarettes per day (19% of all days). This equals one pack of cigarettes per day, which suggests the possibility that such reports might not reflect rounding but actual consumption, structured by the number of cigarettes in a pack. In that case, one might expect better TLFB performance on pack-a-day days. Accordingly, we examined the 599 days on which TLFB reports indicated 20 cigarettes. On these days, EMA reports were especially likely to exceed the TLFB report of 20 cigarettes (39% of days vs 30% of days where TLFB was not equal to 20, p<0.0001). This suggests that reporting of 20 cigarettes on TLFB does not reliably indicate smoking a pack of cigarettes, and suggests that rounded values did not reflect cigarette packaging.

Correlations among sources of self-report

The between-subject correlations among the aggregate CPD estimates were high (Table 2), particularly between Global and pre-monitoring TLFB assessments collected at the same time, before monitoring had started.

Only the EMA and the monitoring TLFB data actually cover the same time period. When averaged across days to obtain a mean for each subject, estimates from EMA and TLFB for the same period correlated β=0.77. This coefficient reflects associations across subjects in their average smoking rate, ignoring day-to-day variability within subjects. In contrast, the within-subject day-by-day association between EMA and TLFB (from hierarchical analysis) was much lower at β=0.29, though still very significant (p<0.0001).

Correspondence with bio-markers of smoking

Salivary cotinine levels averaged 319.47 ng/ml (145.54). We assessed the relationship between cotinine and EMA and TLFB smoking rates in the preceding three days. When

considered separately in univariate models, each measure was significantly associated with cotinine levels, at similar magnitudes (EMA β=0.41; TLFB β=0.38, both p<0.0001). However, when both were entered in a simultaneous multivariate model, only EMA estimates (β=0.33, p<0.01) were uniquely associated with cotinine concentrations; TLFB estimates were not (β=0.10, ns). (The results were similar when we examined Global reports or pre-monitoring TLFB, which were not contemporaneous with the cotinine measure, in place of or in addition to the contemporaneous monitored TLFB.)

CO levels were taken on 3 occasions for most subjects, which afforded an opportunity to test the sensitivity of EMA and TLFB measures to within-subject variations in cigarette consumption. The CO levels averaged 34.90 ppm (14.00), and mean levels did not vary systematically across the three occasions. We examined the relationship of EMA and TLFB with CO readings taken on the same day over the multiple occasions. These analyses controlled for subject effects, and thus essentially model the within-subject variation in CO across occasions on the basis of the within-subject variation in smoking rate across days. Table 3 shows the results. In univariate analyses, the association between EMA and CO was highly significant; in contrast, the association between TLFB and CO was smaller, and not significant (p<.10). When both measures were put in the model, the relationship between EMA and CO remained unchanged, while TLFB had no unique association at all with CO.

The ability of cigarette consumption measures to predict CO within-subject was expected to be greatest among subjects whose CO levels showed substantial variability across measurement occasions, indicating some real changes in smoking. To assess this, we segmented subjects into those with high and low CO variability (SD of CO: High = 15.08 [5.14], n=37; Low= 4.54 [2.59], n=187). For EMA measures, this expected interaction was observed (Table 3): high-CO-variability subjects showed much higher associations between CO and EMA cigarette counts. Significant associations were observed even among low-variability subjects. This was not the case for TLFB-based cigarette consumption measures: the associations were not any stronger among subjects with high CO variability. Furthermore, in each group, the TLFB associations were lower than those estimated for EMA measures, and they were near zero and non-significant in multivariate models including EMA measures.

To address the timeliness of cigarette entries within a day in EMA, we looked at the association between CO measures taken in the clinic, and the number of cigarettes recorded in the preceding 2 hours on each occasion. This "proximal" EMA measure was associated with CO concentrations (β=0.23, p<0.0001), and the effect was significantly greater in subjects with variable CO measures (low variability, β=0.18, p<0.0001; high variability β=0.42, p<0.01; interaction p<0.002). To further isolate the effect of only the most recent smoking, we parsed out the EMA-based smoking rate for the entire day. The association was reduced, but still significant (β=0.14, p<0.001), and again was higher for subjects with more variable COs (low variability, β=0.13, p<0.0002; high variability β=0.21, ns; interaction p<0.005). Thus, EMA entries specifically pick up variation in very recent smoking, suggesting entries were timely.

Discussion

We examined and compared four kinds of self-report of cigarette consumption: Global retrospective estimates, day-by-day retrospective time-line follow-back without contemporaneous self-monitoring, TLFB during self-monitoring, and concurrent recording of cigarettes using an electronic diary to implement EMA methods. The findings suggested that TLFB data collected outside of contemporaneous monitoring add little to Global reports. TLFB reports collected during self-monitoring appeared to be an improvement over

Global reports, and have some validity, but the data also suggested the limits of TLFB, and favored the validity of EMA data, both with regard to digit bias, and also as validated by objective biochemical markers of smoking.

The data illustrated the documented phenomenon of digit bias in Global reports. Further, the data showed that TLFB data collected in the typical way – outside a period of detailed self-monitoring – demonstrate just as much digit bias as Global reports did, with two thirds of entries being even multiples of 10. Further, most subjects filled their TLFB reports for the week with a single uniform value, typically the amount they reported smoking in their Global report. Thus, collecting TLFB data in the absence of self-monitoring seemed to add little to collecting Global reports.

Digit bias was less when TLFB data were collected during a period when subjects were monitoring their smoking and in treatment for cessation, but it was still substantial, with even multiples of 10 over-represented more than 4-fold. (By way of comparison, a United Nations standard considers age data seriously compromised if digit bias or heaping overstates such multiples by a factor of 1.75; United Nations, 2003). Lewis-Esquerre et al. (2005) did not find much digit bias in TLFB reports obtained from light-smoking adolescents, probably because digit bias is minimized when daily consumption is below 10. However, our data suggest that heavier adult smokers do inject substantial digit bias into TLFB reports, which degrades their precision in estimating cigarette consumption, and may diminish their utility for assessing smoking and changes in smoking. In contrast to both Global and TLFB estimates, EMA cigarette counts computed primarily from cigarette-by-cigarette entries throughout the day, were smoothly distributed with no evidence of digit bias or heaping.

Despite the fact that Global and TFB methods showed substantial digit bias, the aggregate estimates of individual smoking rates from the four methods correlated well across subjects – that is, they all similarly discriminate lighter from heavier smokers. Beyond this basic test, however, the measures diverged. On a day-to-day basis during the monitoring period, EMA and contemporaneous TLFB measures only correlated 0.29. Strikingly, this corresponds very closely to recent analyses of day-to-day assessments of pain by EMA and by recall, which found between-subject correlations in the 0.60s through 0.80s, but within-subject correlations of 0.29 (Broderick et al., in press). In both cases, it seems likely that recall is adequate to characterize a person's typical experience, but is stretched beyond its limits when called upon to reconstruct day-by-day experience. The fact that the correlations were so similar in the two settings suggests that this may be a generalizeable phenomenon.

In analyses with biochemical measures, EMA data appeared to perform better as a measure of smoking. Global, TLFB, and EMA measures correlated at roughly comparable levels with individual differences in cotinine levels. However, in multivariate models, retrospective estimates of consumption were no longer associated with cotinine levels, whereas EMA-based estimates retained their relationship to this biochemical marker. In other words, EMA-based measures captured all the between-subject variance in Global or TLFB-based measures, but also captured additional variance in smoking.

The associations between aggregate consumption measures and a single cotinine measure only allowed for assessment of static differences among individuals. Because we had several measures of CO, this measure allowed us to examine how well TLFB and EMA measures captured within-person variation in smoking across days, which each is designed to tap. EMA measures of changes in smoking were validated by concurrent changes in CO levels, but TLFB measures showed no relationship with CO, suggesting that they do not validly tap day-to-day variations in cigarette consumption. The relationship between EMA-assessed

smoking rate and CO was especially robust among smokers whose CO levels (and, presumably, smoking) varied considerably over time, reaching 0.69, lending further support to EMA-based measures.

For analyses that focus on real-time collection of data about the immediate circumstances of smoking, timely recording of cigarettes is important. While we could not objectively confirm when each cigarette was smoked, we indirectly assessed timely recording of smoking by looking at the relationship between CO levels and very recent smoking. The analyses confirmed this temporal relationship between EMA measures and CO. The magnitude of the association was modest, but that was to be expected, given the variability in how much CO people absorb from smoking and how quickly they clear CO from their bodies. The analysis suggests that smokers were recording their cigarettes in a timely way. Substantive findings from EMA monitoring of smoking also seem to validate timely recording of cigarettes. An analysis of the dataset analyzed here found that details of the immediate circumstances surrounding smoking episodes, reported at the time a cigarette was recorded, prospectively predicted subsequent success at smoking cessation (Shiffman et al, 2007). In an independent study, Chandra et al. (2007) found meaningful variations in the distribution of cigarettes by time of day, as recorded by EMA, which also prospectively predicted subsequent success in smoking cessation. Thus, it appears that EMA recording is timely, and is able to capture the timing and circumstances of smoking.

We noted that the various measures of daily consumption produced significantly different estimates. Whereas Global and TLFB measures collected at the outset of the study agreed with each other, the TLFB measures collected during the study were significantly lower, by about two cigarettes per day. This difference might be due to methods differences – specifically, getting TLFB data from a period when smoking was also being monitored cigarette-by-cigarette. However, it is also possible that cigarette consumption actually dropped after enrollment, since subjects were heading towards a quit date and were engaged in treatment. A previous analysis (Shiffman et al., 2002) showed no change in CO levels during the monitoring period, but this does not preclude modest changes in consumption.

EMA estimates of daily smoking rates were about 2.5 cigarettes per day lower than contemporaneous TLFB estimates. One obvious potential explanation is that subjects failed to record some of their cigarettes, whether through neglect, or because some were smoked while subjects were driving or engaged in other activities that precluded timely recording. However, this cannot wholly explain the discrepancy between the two methods, because almost all subjects had days in which they actually recorded more cigarettes than they later thought they had smoked (and it seems unlikely they bothered to make entries when they were not smoking). If reporting shortfalls are not the reason, then why else might recall estimates be higher? Interestingly, a number of analyses of recall suggest that many phenomena – ranging from intensity of pain (Stone, Broderick, Shiffman, & Schwartz, 2004) and craving (Shiffman et al., 2006), to frequency of urination (Homma et al., 2002) and headaches (Van Den Brink, Bandell-Hoekstra, & Abu-Sadd, 2001) – are exaggerated in recall, due to memory biases that make the experience being studied more salient and result in its exaggeration. Cognitive research suggests that people estimate frequency from the cognitive "availability" of memories for the target event (Tversky & Kahneman, 1982; see Shiffman et al., 2008). In this instance, being asked to think about smoking, and perhaps also the experience of having had to record times subjects actually smoked, could make smoking seem more frequent than it is, resulting in higher estimates of consumption. Whereas it is often assumed that smokers are motivated to under-report their smoking, in order to save face, unconscious cognitive biases may actually lead to over-reporting. Further research will be needed to resolve this issue. In the interim, it seems likely that a small proportion of cigarettes may not have been recorded in the EMA data.

These comparisons highlight one of the most important limitations of this study: we had no precise objective measure of smoking, no absolute "Gold Standard" against which to evaluate self-reports. CO and cotinine provided some objective validation, but are not capable of precise cigarette-by-cigarette verification. However, to the extent that they do represent objective markers of smoking, they favored the EMA data over TLFB. Other limitations relate to the study sample: these were relatively heavy smokers who were preparing to quit smoking; the latter could have enhanced reactivity, leading to changes in smoking as a result of self-monitoring (Abrams & Wilson, 1979). However, studies have suggested that EMA monitoring engenders surprisingly little reactivity (Hufford et al., 2002; Shiffman et al., 2008), and analyses of this specific sample showed no systematic decreases in CO over the monitoring period (Shiffman et al 2002). Conversely, the study had some important strengths, including a relatively large sample of smokers monitored over multiple days, with objective demonstration of compliance with active prompting (Shiffman et al., 2002). Further, the EMA recordings provided electronically time-tagged diary records, which are not vulnerable to back-filling (Stone et al., 2002), and the study included multiple biochemical markers of smoking.

The findings suggested that retrospective assessments of daily smoking, even when collected in detail by TLFB, over a relatively short period of recall, and during a period of detailed real-time monitoring, had substantial limitations. TLFB data were subject to substantial digit bias. The fact that TLFB-assessed daily cigarette consumption was not significantly associated with within-subject variations in CO levels questions the validity of this assessment method. On this basis, for studies that focus on cigarette consumption and particularly on within-person changes, TLFB must be regarded as severely limited, even when the recall period is a week or less. In contrast to the findings for TLFB, EMA-based measures showed good associations with biochemical indicators of smoking, and captured both between and within-subject variance in cigarette consumption and biochemical indicators of smoking.

Although this study focused exclusively on cigarette smoking, the analyses may have implications for the use of EMA in other domains. While subject compliance with prompts initiated by an electronic diary can easily be measured, compliance with recording of health-relevant events (e.g., asthma attacks, Hensley et al., 2003; meals,Greeno, Wing, & Shiffman, 2000), at the subject's initiative, is also typically hard to verify. Our findings suggested that compliance with cigarette recording was good, which, if generalizeable, suggests that event recording may be reliable in other EMA domains as well, particularly as recording of cigarettes is a particularly challenging event-monitoring task, because smoking is so frequent and routine. In our sample, subjects recorded individual occasions of smoking an average of more than 20 times per day, and more than 5% of days saw subjects recording over 40 episodes, a heroic burden. In contrast, most EMA event-monitoring protocols ask subjects to monitor events that may occur just a handful of times per day, and may thus be more salient, imposing a lesser burden. In many ways, tracking of cigarette consumption represents the most challenging case for both EMA recording and TLFB recall. Whether retrospective TLFB recall or recall at end of day would adequately capture occurrence of these less frequent, more salient, events is not known. More research is needed to confirm subjects' compliance with event recording, and to assess their ability to accurately recall events. In any case, the present analyses suggested that, for tracking cigarette consumption, TLFB measures were problematic, and EMA recordings were uniquely useful.

Acknowledgments

References

Abrams DB, Wilson GT. Self-monitoring and reactivity in the modification of cigarette smoking. Journal of Consulting and Clinical Psychology 1979;47:243–251. [PubMed: 469070]

Benowitz NL, Jacob P, Ahijevych K, Jarvis MJ, Hall S, LeHouezec J, et al. Biochemical verification of tobacco use and cessation. Nicotine & Tobacco Research 2002;4:149–159. [PubMed: 12028847]

Bradburn NM, Rips LJ, Shevell SK. Answering autobiographical questions: The impact of memory and inference on surveys. Science 1987;236:157–161. [PubMed: 3563494]

Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. Pain. in press.

Chandra S, Shiffman S, Scharf DM, Dang Q, Shadel WG. Daily smoking patterns, their determinants, and implications for quitting. Experimental and Clinical Psychopharmacology 2007;15:67–80. [PubMed: 17295586]

Crawford SL, Johannes CB, Stellato RK. Assessment of digit preference in self-reported year at menopause: Choice of an appropriate reference distribution. American Journal of Epidemiology 2002;156:676–683. [PubMed: 12244037]

Denic S, Khatib F, Saadi H. Quality of age data in patients from developing countries. Journal of Public Health 2004;26(2):168–171. [PubMed: 15284321]

Epstein EE, Labouvie E, McCrady BS, Swingle J, Wern J. Development and validity of drinking pattern classification: binge, episodic, sporadic, and steady drinkers in treatment for alcohol problems. Addictive Behaviors 2004;29(9):1745–1761. [PubMed: 15530719]

Greeno CG, Wing R, Shiffman S. Binge antecedents in obese women with and without Binge Eating Disorder. Journal of Consulting and Clinical Psychology 2000;68:95–102. [PubMed: 10710844]

Hammersley R. A digest of memory phenomena for addiction research. Addiction 1994;89(3):283–293. [PubMed: 8173494]

Hatsukami D, Mooney M, Murphy S, LeSage M, Babb D, Hecht S. Effects of high dose transdermal nicotine replacement in cigarette smokers. Pharmacology, Biochemistry, and Behavior 2007;86(1):132–139.

Hensley MJ, Chalmers A, Clover K, Gibson PG, Toneguzzi R, Lewis PR. Symptoms of asthma: Comparison of a parent-completed retrospective questionnaire with a prospective daily symptom diary. Pediatric Pulmonology 2003;36:509–513. [PubMed: 14618643]

Hessel PA. Terminal digit preference in blood pressure measurements: Effects on epidemiological associations. International Journal of Epidemiology 1986;15:122–125. [PubMed: 3957533]

Homma Y, Ando T, Yoshida M, Kageyama S, Takei M, Kimoto K, et al. Voiding and incontinence frequencies: variability of diary data and required diary length. Neurourology and Urodynamics 2002;21:204–209. [PubMed: 11948713]

Hufford MR, Shields AL. Electronic diaries: an examination of applicationsand what works in the field. Applied Clinical Trials 2002;11:46–56.

Hufford MR, Shields AL, Shiffman S, Paty J, Balabanis M. Reactivity to ecological momentary assessment: an example using undergraduate problem drinkers. Psychology of Addictive Behavior 2002;16:205–211.

Hughes JR, Carpenter MJ. Does smoking reduction increase future cessation and decrease disease risk?A qualitative review. Nicotine &Tobacco Research 2006;8(6):739–749. [PubMed: 17132521]

Klesges RC, Debon M, Ray JW. Are self-reports of smoking rate biased? Evidence from the second National Health and Nutrition Examination Survey. Journal of Clinical Epidemiology 1995;48:1225–1233. [PubMed: 7561984]

Lewis-Esquerre JM, Colby SM, Tevyaw TO, Eaton CA, Kahler CW, Monti PM. Validation of the timeline follow-back in the assessment of adolescent smoking. Drug and Alcohol Dependence 2005;79(1):33–43. [PubMed: 15943942]

Miller WR, DelBoca FK. Measurement of drinking behavior using the Form 90 family of instruments. Journal of Studies on Alcohol 1994;12 Suppl.:112–118.

Monti PM, Barnett NP, Colby SM, Gwaltney CJ, Spirito A, Rohsenow DJ, et al. Motivational interviewing versus feedback only in emergency care for young adult problem drinking. Addiction 2007;102:1234–1243. [PubMed: 17565560]

Nagi MH, Stockwell EG, Snavley LM. Digit preference and avoidance in the age statistics of some recent African censuses: Some patterns and correlates. International Statistical Review 1973;41(2): 161–174.

Rowland ML. Self-reported weight and height. American Journal of Clinical Nutrition 1990;52:1125–1133. [PubMed: 2239790]

Shiffman S, Balabanis MH, Gwaltney CJ, Paty JA, Gnys M, Kassel JD, Hickcox M, Paton SM. Prediction of lapse from associations between smoking and situational antecedents assessed by ecological momentary assessment. Drug and Alcohol Dependence 2007;91:159–168. [PubMed: 17628353]

Shiffman S, Ferguson SG, Gwaltney CJ, Balabanis MH, Shadel WG. Reduction of abstinence-induced withdrawal and craving using nicotine replacement therapy. Psychopharmacology 2006;184:637–644. [PubMed: 16261317]

Shiffman S, Gwaltney CJ, Balabanis M, Liu KS, Paty JA, Kassel JD, Hickcox M, Gnys M. Immediate antecedents of cigarette smoking: An analysis from ecological momentary assessment. Journal of Abnormal Psychology 2002;111:531–545. [PubMed: 12428767]

Shiffman S, Paty JA, Gwaltney CJ, Dang Q. Immediate antecedents of cigarette smoking: An analysis of unrestricted smoking patterns. Journal of Abnormal Psychology 2004;113(1):116–171. [PubMed: 14992664]

Shiffman S, Stone AA, Hufford M. Ecological momentary assessment. Annual Review of Clinical Psychology 2008;4:1–32.

Shryock, HS.; Siegel, JS. Methods and Materials of Demography. New York: Academic Press; 1976.

Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. British Medical Journal 2002;324:1193–1194. [PubMed: 12016186]

Stone AA, Broderick JE, Shiffman S, Schwartz JE. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. Pain 2004;107:61–69. [PubMed: 14715390]

Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Annals of Behavioral Medicine 1994;16:199–202.

Swan GE, Habina K, Means B, Jobe JB, Esposito JL. Saliva cotinine and recent smoking- Evidence for a nonlinear relationship. Public Health Reports 1993;108(6):779–783. [PubMed: 8265764]

Toll BA, Cooney NL, McKee SA, O'Malley SS. Do daily interactive voice response reports of smoking behavior correspond with retrospective reports? Psychology of Addictive Behaviors 2005;19(3):291–295. [PubMed: 16187808]

Tversky, A.; Kahneman, D. Availability: A heuristic for judging frequency and probability. In: Kahneman, D.; Slovic, P., editors. Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press; 1982. p. 163-178.

United Nations. Demographic Yearbook special census topics: Basic population characteristics. New York: United Nations; 2003.

Van Den Brink M, Bandell-Hoekstra E, Abu-Sadd H. The occurrence of recall bias in pediatric headache: a comparison of questionnaire and diary data. Headache 2001;41:11–20. [PubMed: 11168599]
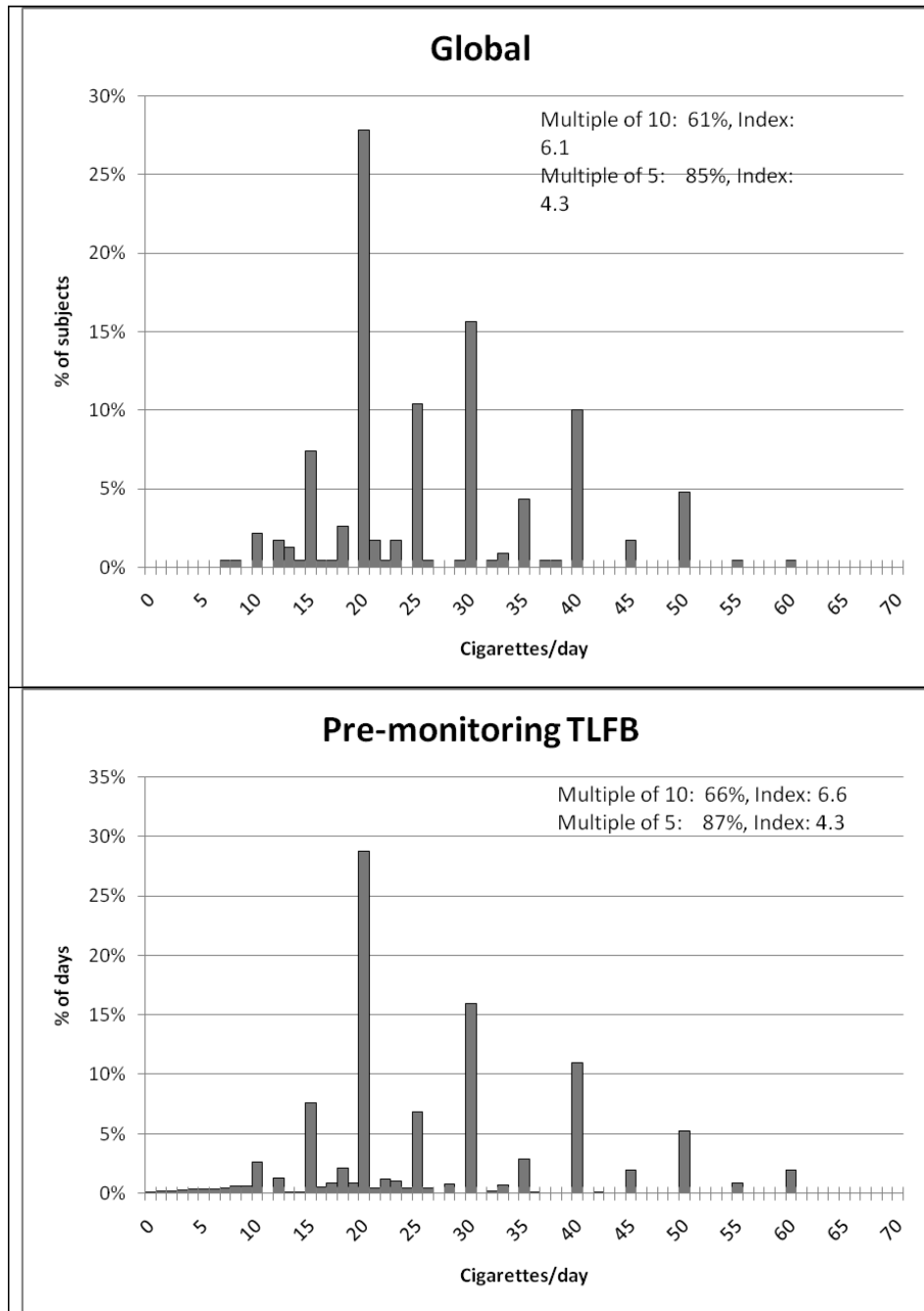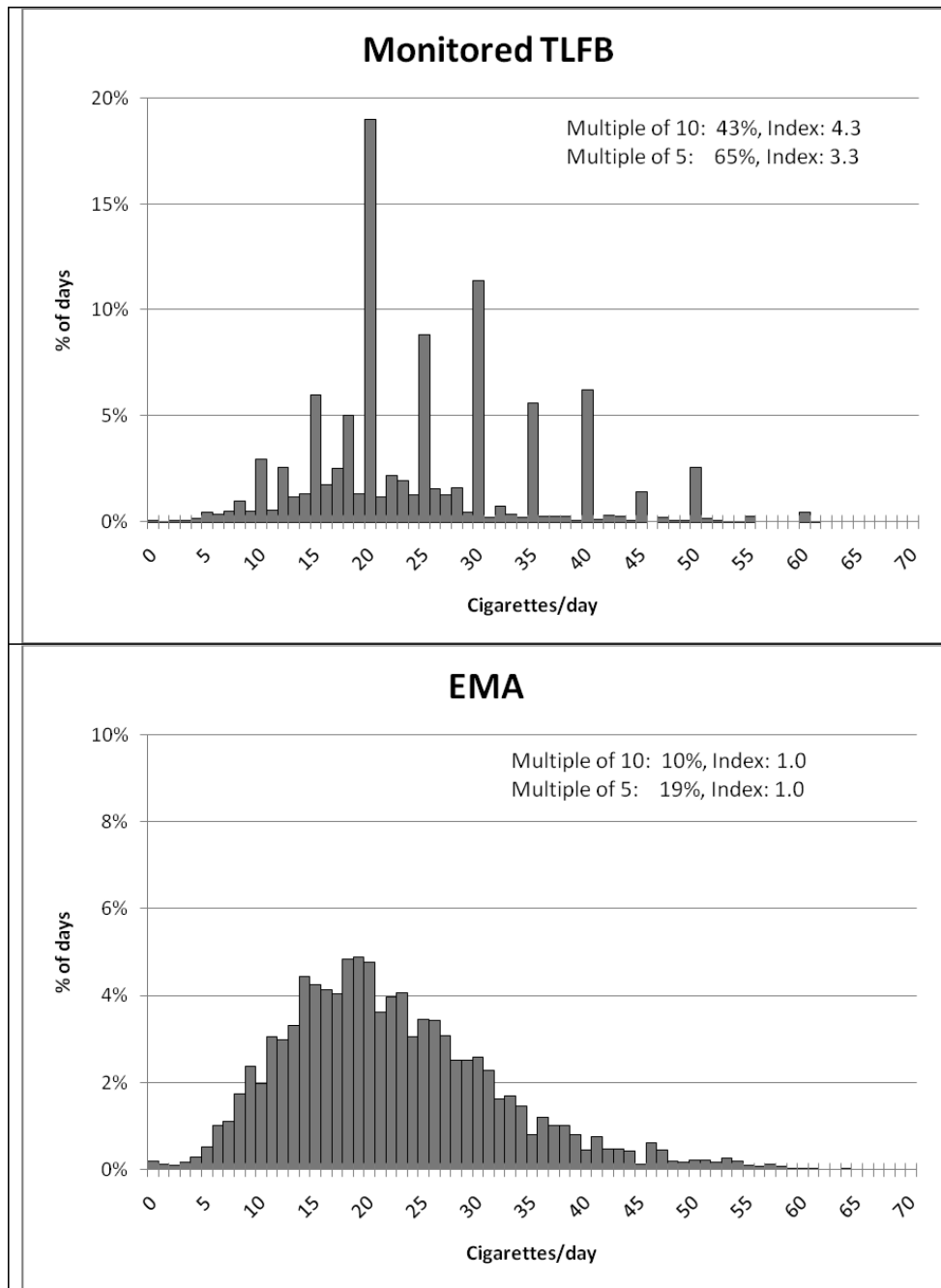
Figure 1.
The distribution of daily cigarette consumption, as reported by Global self-report (a single value for each subject), pre-monitoring TLFB (for the 7 days prior to entry into the study), Monitored TLFB (14 days during EMA monitoring), and EMA recording. The last 3 are based on estimated cigarette consumption for individual days. The figures also show the percentage of days where the values were even multiples of 10 and 5, and the value of Whipple's index, which expresses that percentage as a function of the percentage expected.

NIH-PA Author Manuscript

Table 1

Characteristics of daily cigarette consumption, estimated by four methods

	Daily Consumption		Rounded at 10		Invariant
	Mean	**SD**	**Mean**	**SD**	**%**
Global report	26.36[a]	10.35	*60.9%	48.9%	--
TLFB (pre-monitoring)[1]	26.66[a]	12.83	*64.3%[a]	40.3%	51.1% [a]
TLFB (monitored)	24.61[b]	9.35	*42.8%[b]	29.6%	3.5% [b]
EMA	21.97[c]	8.56	10.2%[c]	8.3%	0.0%[c]

Note. Rounded at 10 refers to the percent of days on which reported consumption was an even multiple of 10; it is equal to 10 x Whipple's index. Invariant refers to the percent of subjects who reported exactly the same cigarette consumption on all days. Daily Consumption means with different superscript letters differ significantly from each other, p<0.0001. Round 10 means with different superscript letters differ significantly from each other, p<0.0003; Global report could not be tested, as it was a single dichotomous value (rounded or not) for each subject (whereas the other values are the percentage of days within a subject that were rounded). Invariant percentages with different subscripts differ significantly from each other, p<0.0001.

*
significantly different from the proportion expected under a random distribution, p<0.0001

[1]Subset of subjects, n=194

Table 2

Between-subject correlations among daily cigarette consumption measures

	Global	Pre-monitoring TLFB	Monitored TLFB
Global	--		
Pre-monitoring TLFB	0.94	--	
Monitored TLFB	0.85	0.82	--
EMA	0.77	0.73	0.87

All correlations were significant, p<0.0001

Table 3

Associations between daily cigarette consumption and CO over multiple days

| | All subjects (n=230) | Variability in CO across occasions | | |
		Low (n=137)	High (n=87)	Interaction effect
<u>Univariate</u>				
EMA	**** 0.35	**** 0.24	*** 0.69	****
TLFB	0.12	0.12	0.12	
<u>Multivariate</u>				
EMA	**** 0.34	**** 0.23	*** 0.72	****
TLFB	0.01	0.05	−0.11	

*
 $p<0.05$;

**
 $p<0.01$;

***
 $p<0.001$;

****
 $p<0.0001$

Entries are standardized regression coefficients (β), which are equivalent to the correlation coefficient r in the univariate case.