Compliance With Saliva Sampling Protocols: Electronic Monitoring Reveals Invalid Cortisol Daytime Profiles in Noncompliant Subjects

BRIGITTE M. KUDIELKA, PhD, JOAN E. BRODERICK, PhD, AND CLEMENS KIRSCHBAUM, PhD

Objective: Ambulatory saliva collections for subsequent analysis of free cortisol levels are now frequently applied to measure adrenocortical activity in healthy subjects and patient populations. Despite the prime importance of accurate timing of saliva collection outside the laboratory, no data are available on the compliance of study participants following a given sampling protocol. This study investigated how accurately subjects adhered to the instructions to collect six saliva samples throughout 1 day. Methods: Subjects were instructed to collect six saliva samples throughout 1 day (directly after awakening, 30 minutes after awakening, 11 AM, 3 PM, 8 PM, 10 PM). Objective compliance was measured using an electronic monitoring device given to the subjects either with (“informed” \( N = 23 \)) or without (“noninformed” \( N = 24 \)) their knowledge of the nature of the device. Data on subjective compliance were obtained by self-report. Results: Thirty-one subjects (74%) were found to comply with the sampling instructions, and 11 (26%) failed at least once to obtain the saliva sample at the correct time of day. Nine of the 11 noncompliant subjects (82%) had two or more noncompliant samples. Fifty-five percent (6 of 11) of the noncompliant subjects took sample 2 outside the sampling window. The circadian cortisol profile differed significantly between compliant and noncompliant subjects (\( F = 7.98, p = .007 \)). The most important effect of compliance was seen in the rise of cortisol at awakening. Compliant subjects showed a robust increase, whereas noncompliant individuals had only minimal changes from baseline at 30 minutes after awakening (\( t = 2.89, p = .007 \)). Thus the steepness of the circadian cortisol decline was greater for compliant subjects (\( t = 2.10, p = .043 \)). Furthermore, the informed group adhered more closely to the sampling protocol than the noninformed subjects (\( p = .001 \)). Self-reported compliance also differed significantly between study groups (\( p = .03 \)). In the noninformed group, self-reported sampling accuracy was significantly higher than objectively measured compliance (\( p = .03 \)); the two measures were similar in the informed group (\( p = \) NS). Conclusions: A significant number of subjects did not obtain saliva samples reliably in an ambulatory setting. This can partially invalidate the cortisol results and mask potential differences between subject groups of interest. We therefore recommend the use of electronic monitoring devices or other suitable methods and that study participants be informed about the device when ambulatory saliva collection is performed. Key words: electronic monitoring, compliance, sampling accuracy, salivary cortisol, awakening cortisol response, circadian rhythm.

ANOVA = analysis of variance; SEM = standard error of mean.

INTRODUCTION

Patient compliance with a prescribed drug regimen has been an important issue in clinical pharmacology, but measurement of compliance has presented challenges. For example, compliance has been estimated by counting returned, unused medications (tablet counts); by interviewing patients; or by having patients complete diaries and questionnaires. These methods are strongly biased toward overestimation of compliance due to social pressure for the patient to seem compliant to the physician (1–9). Since 1977 various methods of electronic medication event monitoring have been used to assess drug usage (10–14), and two chemical markers, digoxin and phenobarbitone, have been introduced to assess compliance biochemically (15–17). Although these chemical markers verify drug ingestion, they give no information about the time of drug ingestion. In contrast, electronic monitoring devices provide detailed date and time records of opening the medication container, but they cannot prove that medication was ingested. Consequently, the combination of these two methods, electronic monitoring in combination with chemical markers, is currently regarded as the “gold standard” for compliance measurement in clinical pharmacology.

Some evidence from clinical observations indicates that simply informing patients that their dosing is being recorded does not enhance patient adherence (18, 19). Other studies suggest that compliance can be improved in some patients by reviewing monitoring data with them (2). For example, in the Lung Health Study, it was shown that patients who received feedback about their dosing times showed significantly better compliance with the prescribed regimen compared with a control group that did not receive feedback (3, 20, 21). Kruse and Weber (22) reported that patients

From the Department of Experimental Psychology II (B.M.K., C.K.), University of Düsseldorf, Germany; and the Department of Psychiatry and Behavioral Science (J.E.B.), Applied Behavioral Medicine Research Institute, State University of New York at Stony Brook, Stony Brook, New York.

Address reprint requests to: Clemens Kirschbaum, PhD, Department of Experimental Psychology II, University of Düsseldorf, Univer-

sitätstr. 1, 40225 Düsseldorf, Germany. Email: ck@uni-duesseldorf.de

Received for publication July 27, 2001; revision received March 5, 2002.

DOI: 10.1097/01.PSY.0000058374.50240.BF
who received information about being electronically monitored adhered to the prescribed drug protocol more closely than noninformed patients. Different patient groups, however, were not randomly assigned to the experimental groups (informed vs. noninformed).

Research involving multiple daily entries of symptom reports is another arena in which subject compliance with a sampling protocol has been questioned. Eliminating retrospective reporting bias is achieved when subjects provide symptom reports at the designated sample time. Self-reported compliance with written diary protocols has been high (23–25). However, recent work has shown that actual compliance among patients reporting pain in a paper diary is exceedingly low despite self-reports of high compliance (26).

Although some data on patient compliance with drug regimens from electronic monitoring and symptom diaries are already available (20, 27–33), there is no comparable data concerning compliance with (or the accuracy of) saliva sampling. In psychobiological research, many study designs rely on the honesty and motivation of each individual subject to provide protocol-adherent data. Because of the many advantages of saliva sampling as compared with blood sampling (34, 35), an increasing number of studies are using ambulatory saliva collection for measurement of hormones in saliva. The standard methodology includes detailed instructions about the purpose of saliva collection, the importance of exact timing of the samples, the specific procedure for taking and storing the samples, and cautions about eating and brushing teeth before the sample. Subjects are provided with all of the materials needed for saliva collection and sent into their natural environment to gather the samples at the designated times. This standard methodology was examined in this study.

Among the various analytes that can be measured in saliva today, free cortisol is currently the hormone most frequently measured. The ease of sampling is especially attractive for large-scale research projects, some of which include more than 20,000 samples (eg, see Ref. 36). Accurate saliva sampling is vital to the assessment and interpretation of cortisol profiles because cortisol shows a typical circadian rhythm with up to 100% change within 30 minutes (eg, after awakening, meals, stress; eg, see Refs. 34 and 35).

Despite the widespread use of ambulatory saliva collection in an increasing number of studies, it is still unknown how closely subjects follow the prescribed saliva sampling protocol in their natural environment. This study investigated objective compliance with a collection schedule and the effects of noncompliance on circadian cortisol profiles. The study also examined the possible impact on compliance of awareness by subjects that sampling times were being electronically monitored.

METHODS

Subjects

Forty-seven community-dwelling subjects aged 15 to 75 years (mean age, 30.9 ± 1.97 years) volunteered to collect six saliva samples throughout 1 day in their natural environment. The study sample consisted of 28 women and 19 men. Thirty-two subjects stated that they were nonsmokers, and 15 subjects indicated that they smoked 4 to 20 cigarettes per day. Thirty-eight subjects were apparently healthy, and nine participants reported minor health problems.

Study Protocol

The study protocol followed an experimental design. Twenty-three subjects were randomly chosen to be informed about the electronic monitoring device and procedure (informed group). The remaining 24 subjects were not informed (noninformed group).

Salivette sampling devices (Sarstedt, Rommelsdorf, Germany) were used for saliva collections. Six saliva sampling swabs were removed from the original plastic tubes by the experimenter and put in an electronic drug exposure monitor (eDEM, Aardex Ltd., Switzerland). All subjects received the same detailed instructions describing the Salivette saliva sampling method and stressing the importance of exact timing of samples. Subjects were told to obtain six saliva samples during a single day: directly after awakening, 30 minutes later, and at 11 AM, 3 PM, 8 PM, and 10 PM. Subjects were instructed not to brush their teeth before completing saliva sampling to avoid contamination of saliva with blood caused by microinjuries to the oral cavity. Smoking, eating, and drinking beverages containing alcohol, caffeine, or fruit juices were not allowed for 30 minutes before saliva sampling or during the saliva collection. Besides these restrictions, subjects were otherwise free to follow their normal daily routines on the sampling day. They were instructed to open the plastic container (ie, electronic monitor) and remove a single cotton roll at each designated sampling time. Subjects were told that the swabs must remain in the container until usage to ensure valid hormone analysis in the laboratory. After saliva collection, each wet swab was stored in a plastic Salivette tube labeled with the designated sampling time by the experimenter.

In addition, all participants filled out a questionnaire on sociodemographic details (gender, smoking habits, and health problems). Moreover, they noted the exact time of sampling for each saliva sample in a timetable. This record was used to assess self-reported compliance. Subjects did not receive reimbursement for study participation. The study protocol was approved by the University of Düsseldorf ethics committee.

Cortisol Analysis

Salivary cortisol was assayed with a time-resolved immunoflourescence detection (Delfia) as described by Dressendorfer et al. (37). Intraassay and interassay variability were below 12%.

Statistical Analysis

A special interface and the software program PowerView (version 1.4.0) were used to transfer data from the electronic monitor to
COMPLIANCE WITH SALIVA SAMPLING

Deviations From Sampling Protocol

Across all subjects, a total of 24 of 252 samples (10%) were taken outside the compliance windows. In the first set of analyses, objective and subjective sampling compliance were compared between the two experimental groups. The electronic monitor data revealed that the total group showed a mean deviation of 13 ± 4 minutes per sample (range, 0–264 minutes) from the target sampling time. To compare time deviations between the experimental groups, minutes of deviation were summed across the six samples. Results showed that the informed group had a mean cumulated deviation of 30 ± 7 minutes compared with 125 ± 26 minutes in the noninformed group ($t(23.9) = 3.6, p = .001$; Table 1). Analyses of the self-report data showed that the informed group reported a mean deviation of 25 ± 8 minutes from the target sampling times compared with 68 ± 18 minutes in the noninformed group ($t(28.5) = 2.3, p = .03$; Table 1). A repeated-measures ANOVA yielded a significant interaction of group by compliance (informed/noninformed vs. objective/subjective deviation; $F(1,40) = 4.87, p = .033$). t tests for dependent samples showed that objective and subjective compliance differed significantly in the noninformed group ($t(20) = -2.4, p = .03$), but there was no difference between objective and subjective data in the informed group ($t(20) = -1.3, p = .22$; Table 1). Ninety-seven percent of samples from informed subjects were compliant according to the criteria outlined in “Methods”; 84% of samples from noninformed subjects were compliant.

Impact of Compliance on Cortisol Profiles

In the second set of analyses, cortisol profiles were compared between compliant and noncompliant participants. Subjects were coded as compliant or noncompliant according to the criteria noted in “Methods.” Complete cortisol profiles (all six samples) were available for 42 participants. Five participants did not provide enough saliva in at least one sample or were missing one or more of the samples to plot the morning cortisol profile; data for these participants were excluded from this analysis. Among the 42 subjects included, 31 subjects (74%) fulfilled the defined compliance criteria; 11 subjects (26%) did not collect the six samples within the designated time windows. Eighty-two percent (9 of 11) of the noncompliant subjects had two or more noncompliant samples. Fifty-five percent (6 of 11) of the noncompliant subjects took sample 2 outside the sampling window. Among the 24

| TABLE 1. Differences Between Scheduled Sampling Times: Subjective (Self-Reports) and Objective Compliance Data (Electronic Monitoring) in the Informed and Noninformed Groups |
|-----------------|-----------------|-----------------|
| Deviation (min) | Informed Group  | Noninformed Group |
| Subjective compliance | 25 ± 8          | 68 ± 18         | $p < .03$ |
| Objective compliance  | 30 ± 7          | 125 ± 26        | $p = .001$ |

a Deviations are summed over six sampling times. Values are mean ± SEM.
noncompliant samples, 46% (11) were for the early morning rise (samples 1–2), and the other 54% (13) were for the remainder of the day (samples 3–6).

As shown in Table 2, the two groups did not differ with respect to age, gender, number of smoked cigarettes, or awakening time, but they differed significantly in the composition of subjects from the informed and noninformed groups (Fisher exact test, \( p = .035 \)).

A repeated-measures ANOVA of the circadian cortisol profiles revealed a significant difference between the compliant and noncompliant groups (\( F(1,40) = 7.98, p = .007 \)). Although showing similar waking cortisol levels, compliant subjects had a much greater net increase within the first 30 minutes after awakening (11.03 vs. 3.31 nmol/liter) than noncompliant subjects (\( t = 2.87, p = .007; \) Fig. 1). Furthermore, a Pearson correlation resulted in a significant negative relationship between the time deviation in sample 2 (30 minutes) and the cortisol increase (\( r = -0.32, p = .037 \)), indicating that the larger the time deviation, the smaller the observed cortisol increase. Because the measured circadian peak cortisol levels have a major impact on the observed steepness of the cortisol decline over the remainder of the day, noncompliant subjects also showed a relatively smaller maximum–minimum difference in cortisol values than compliant participants (\( t = 2.10, p = .043 \)).

**DISCUSSION**

This is the first study that measured compliance with saliva sampling in an ambulatory setting. Because collection of saliva samples for endocrine measures is becoming an increasingly important research tool in psychosomatic medicine and other disciplines, it is crucial to investigate how closely subjects follow a given sampling scheme. Special precautions should be taken with regard to exact timing of sample collection when measuring hormones known to have a profound circadian rhythm (like cortisol as opposed to dehydroepiandrosterone-sulfate, for example). Although salivary cortisol is now measured in increasing frequency by many laboratories, no data have been available thus far on collection compliance and the possible impact on the circadian cortisol rhythm.

Using an electronic monitoring device that has been used in clinical pharmacology studies for a number of years, we observed that 26% of subjects did not follow the requested sampling protocol for at least one sample; 21% did not follow the protocol for two or more of the six samples. Furthermore, more than half of the noncompliant samples were obtained during the early morning rise, a critical sample for characterizing an individual’s circadian profile. Participants showed a mean time deviation of 13 minutes per sample, but the time deviations were sometimes striking, ranging up to 264 minutes. Because we allowed rather large time windows for compliance with the target sampling times (ie, sampling time ± 60 minutes for the samples to be obtained at 11 AM, 3 PM, 8 PM, and 10 PM), it was quite surprising that one of four subjects did not collect the samples within the liberal time windows. The rather low compliance observed was disappointing given that the subjects were asked to collect only six samples throughout a single day. Because the reliability of results increases with repeated measurements, more studies now include salivary cortisol measurements over several days (eg, see Refs. 41–46). It is reasonable to speculate that even lower compliance would be observed with increased collection load.

This study provides further evidence that researchers cannot rely on participants’ self-reports of sam-

---

**TABLE 2. Differences in Demographic Variables in Compliant vs. Noncompliant Subjects and Distribution of Informed vs. Noninformed Subjects (means ± SEM)**

<table>
<thead>
<tr>
<th></th>
<th>Compliant Group ((N = 31))</th>
<th>Noncompliant Group ((N = 11))</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>30.03 ± 2.72</td>
<td>31.18 ± 2.34</td>
<td>.75</td>
</tr>
<tr>
<td>Men ((N))</td>
<td>10</td>
<td>6</td>
<td>.19</td>
</tr>
<tr>
<td>Women ((N))</td>
<td>21</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Cigarettes per day</td>
<td>2.45 ± 0.81</td>
<td>7.73 ± 2.74</td>
<td>.21</td>
</tr>
<tr>
<td>Awakening time (h:min)</td>
<td>7:33 ± 0:19</td>
<td>7:32 ± 0:33</td>
<td>.98</td>
</tr>
<tr>
<td>Informed ((N))</td>
<td>18</td>
<td>2</td>
<td>.035</td>
</tr>
<tr>
<td>Noninformed ((N))</td>
<td>13</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

---

**Fig. 1.** Morning and day salivary cortisol profiles (nmol/liter; mean ± SEM) in compliant vs. noncompliant subjects.
COMPLIANCE WITH SALIVA SAMPLING

pling times. Whereas our subjects who were unaware of the monitoring indicated that they deviated from the protocol timing by a total of 68 minutes, actual deviation based on the electronic monitor data was twice that. This probably results from participants wishing to please the researcher and provide “good” data rather than outright deceit. Informing participants that their sampling behavior was being electronically monitored was effective in producing greater compliance and increased veracity of subjects’ self-reports.

The effects of low compliance on saliva collection can be quite dramatic if the concentration of the analyte is expected to change rapidly over time. As shown here with a single-day cortisol profile, noncompliance can lead to erroneous interpretations. In this study, noncompliant subjects showed a much smaller cortisol increase after awakening in the morning. Thirty minutes after awakening, free cortisol levels show a significant increase, normally resulting in the daytime peak values of the circadian rhythm, with continuously decreasing concentrations afterward (35, 47). In several independent studies, the authors observed that about 75% of all subjects show a significant cortisol response to awakening that typically peaks at 30 minutes after waking (eg, Ref. 47). Although the latency of the awakening response peak will differ between subjects and therefore will add noise to the data (independent of compliance), the observed results are highly significant. These results suggest that precise sampling of the second sample (30 minutes after awakening) is very important to gain a valid measure of the cortisol awakening response. We found that the larger the time deviation for sample 2, the smaller the observed awakening cortisol increase. If subjects delay sample 2, they obviously miss the peak, and the resulting awakening increase turns out to be smaller.

Without information on compliance, such an awakening cortisol pattern might have been misinterpreted as “blunted” or indicative of a sluggish adrenocortical responsiveness. Likewise, had we searched for individuals with a “flat” circadian rhythm, the noncompliant subjects in the present study might have been flagged for subsequent in-depth investigation. Flat circadian rhythms have recently garnered attention since Sephton et al. (48) found a flattened cycle to be predictive of early mortality in a group of former breast cancer patients. Spiegel et al. (49) reported that sleep debt was associated with increased cortisol levels in the afternoon and with poorer glucose tolerance. Thus, a flat circadian cortisol rhythm could reflect unfavorable glucose metabolism resulting in a host of potential health problems. In a reanalysis of four studies, Stone et al. (50) found supportive evidence for the previously published notion (51) that in samples of healthy adults, about 10% to 15% have relatively flat circadian cortisol profiles during the waking hours. Because no data on compliance are available for these subjects, it is impossible to test whether the flat rhythms are a true characteristic or rather a result of poor compliance. Nevertheless in this study, even in compliant subjects (according to the electronic monitor), there were individuals with rather unusual cortisol patterns. In other words, the present results must not be understood as suggesting that flat cortisol rhythms measured in an ambulatory setting are simply reflecting noncompliance with the sampling protocol.

The present data are consistent with the growing evidence that research participants are resistant to adherence with sampling or treatment protocols in the natural environment. Whether it is taking medication in the prescribed manner, completing daily ratings of symptoms and mood, or collecting six saliva samples in 1 day, people perform poorly when left on their own. It is noteworthy that even with an overall sampling compliance of 84% by noninformed subjects, a rate that might be viewed as acceptable, it resulted in deviant cortisol profiles and potentially invalid conclusions. Some researchers have suspected problems and have sought to resolve them by signaling subjects with an alarm on a watch at each sampling time (52, 53). Although this is a logical solution, there are no objective data to support the efficacy of this approach. Again, researchers have relied on the self-reports of subjects about their compliance. Work recently completed in one of our laboratories has documented that compliance is only marginally improved with signals from watches for paper diary entries, and it remains severely low despite high levels of self-reported compliance (J. E. Broderick, personal communication, October 30, 2001). Electronic compliance monitoring of informed subjects raises compliance to more acceptable levels, but even then deviations from protocol are evident and acknowledged by subjects (20, 25, 54).

Thus, it would seem that unless patients know that the researcher will systematically detect noncompliance, motivation to adhere to the protocol will often be insufficient.

It should be noted that besides the circadian rhythm, cortisol levels are influenced by several other factors (eg, stressful events, physical exercise, food intake, acute nicotine consumption). Such influences can affect what would otherwise be a typical cortisol profile; for example, stress and exercise can produce a spike in cortisol. Apart from sampling times, other types of noncompliance with the study protocol include poor adherence to the smoking, food, and beverage restrictions, each of which can change a cortisol...
reading. Studies are needed to clarify the impact of all forms of noncompliance for saliva sampling.

In summary, we present evidence to suggest that even with a relatively low sampling burden, a significant number of individuals will not fully comply with an ambulatory saliva sampling regimen unless they are aware of being monitored electronically. Although it is possible that other strategies might also improve compliance, we suggest use of electronic monitoring in studies collecting saliva in an ambulatory setting. Electronic monitoring substantially improves adherence to a sampling protocol, and it allows the researcher to identify samples collected outside the protocol window to eliminate them if necessary. We can presume that the ability to eliminate or otherwise reclassify samples taken at the wrong times will reduce error variance in studies. We can only speculate at this point how much improvement in sampling compliance will increase the reproducibility and validity of such studies. At least some of the unexpected or atypical data on salivary cortisol in the past can be attributed to compliance problems.

This study was supported by the European Science Foundation Program on Social Variations in Health Expectancy in Europe. We thank Uta Döll, Thorsten Günther, Vera Klinkenberg, Vera Mertens, Yvonne Mühlig, Kerstin Polmans, Christina Rohbeck, and Ramya Uthayakumar for recruitment of study participants and data processing. We also thank Clemens G. Hadstein for creation of the applied Excel program.

REFERENCES

COMPLIANCE WITH SALIVA SAMPLING


