Use of salivary biomarkers in biobehavioral research: cotton-based sample collection methods can interfere with salivary immunoassay results

Elizabeth A. Shirtcliff a, Douglas A. Granger a,*, Eve Schwartz b, Mary J. Curran b

a Department of Biobehavioral Health, Behavioral Endocrinology Laboratory, The Pennsylvania State University, 315 East Health and Human Development, University Park, PA 16802-6509, USA
b Salimetrics, LLC, State College, PA, USA

Received 11 April 2000; accepted 25 July 2000

Abstract

In a series of studies, we evaluated the susceptibility of immunoassays for saliva biomarkers to interference effects caused by cotton materials used to absorb saliva during sample collection. Salivary assay results for testosterone, DHEA, progesterone, and estradiol are artificially high, and for sIgA artificially low, when samples are collected using cotton absorbent materials. In contrast, results for salivary cortisol, DHEA–S, and cotinine are not affected by the use of cotton collection methods. The order of individual results from samples collected using cotton versus no-cotton methods for certain markers is not conserved, suggesting that for some biomarkers this collection method can be a significant source of unsystematic error. It was shown, using DHEA as an example, that the cotton interference effect is of sufficient magnitude to attenuate the association between serum and saliva levels. Awareness of this issue is critical to ensure measurement validity in future studies and analyses of archived samples collected using cotton materials. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Saliva; Hormones; Cotinine; sIgA; Immunoassay; Sample collection
1. Introduction

The integration of salivary biomarkers into behavioral and health-related research has increased exponentially (see reviews by Kirschbaum et al., 1992; Malamud and Tabak, 1993). Monitoring biomarkers in saliva has several distinct advantages over doing so in other biological fluids (i.e., urine, serum or plasma). Saliva sampling is relatively non-invasive, enabling collection of samples in special populations (e.g., infants, children, elderly), and in many circumstances in which blood or urine sampling is not viable. Salivary levels accurately reflect the unbound, biologically active, fraction of many serological markers in the general circulation (Riad-Fahmy et al., 1984; Vining et al., 1983). Also, unless visibly contaminated with blood, human saliva is not considered a class II biohazard (US Center for Disease Control), affording researchers and institutions administrative and safety benefits.

As the application of salivary measures in research has increased, so too has our understanding of circumstances and conditions that may influence the validity of these assessments. Studies warn that materials placed in subjects’ mouths to stimulate saliva flow (e.g., powdered drink mix crystals containing citric acid) have the potential to lower sample pH (<3.0) and affect immunoassay results by interfering with antibody binding (Schwartz et al., 1998). Given the 10-to-100:1 concentration difference between blood and saliva levels of most serological markers, atypical blood or serum leakage into the oral mucosa (i.e., due to periodontal disease, teething, shedding teeth, injury, or vigorous cleaning) can affect the integrity of quantitative estimates of salivary concentrations (e.g., Granger, 2000; Hertsgaard et al., 1992). Moreover, studies with infants caution that saliva can be contaminated by hormone-like substances present in breast milk (Magnano et al., 1989). Food products may also contain substances (e.g., bovine hormones in dairy products) with the potential to cross-react with antibodies used in salivary immunoassays. In this brief report, we describe another factor with potential to compromise quantitative measurement of biomarkers in saliva.

The most commonly used method of saliva collection involves the use of cotton-based materials to absorb saliva from subjects’ mouths. Several commercially available collection devices use this approach and many research groups use sterile cotton dental rolls or cotton-tipped applicators to absorb sample. Saliva is expressed from the saturated cotton using, for instance, a 10 cc needleless syringe or by centrifugation. A few studies warn that use of such absorbent materials can induce artificially high results in salivary immunoassays for testosterone (Dabbs, 1991; Granger et al., 1999a), DHEA (Granger et al., 1999b), and estradiol (Shirtcliff et al., 2000). As the range of salivary markers available to be included in biobehavioral studies increases, and as samples are archived for testing in the future, it is critical that investigators know which markers and assays are, and are not, influenced by this common collection method.

The purpose of this study is to underscore the potential of cotton-based absorbent materials to affect quantitative estimates of the concentration of eight common biological markers assayed in saliva. We estimate the impact of this collection method on marker concentration level as well as the rank of individual scores within each
biomarker’s distribution. A secondary aim is to determine whether the magnitude of this effect is similar or different for radioimmuno- or enzyme-immunoassay protocols. Finally, we use DHEA as an example to illustrate that immunoassay results from samples collected using cotton to absorb saliva may underestimate the levels present in the general circulation.

2. Method

2.1. Sample collection and preparation

Saliva was sampled from several independent groups (n ranging from 4 to 21) of young adults. Subjects rinsed their mouths with water, waited five minutes, then expectorated 6–10 mls of saliva through a short plastic straw into a collection vial. Samples from each subject were then either left untreated or passed through a sterile cotton dental roll (Richmond, Charlotte, NC), and/or filtered through the cotton swab used in the plain cotton version of the Salivette device (Sarstedt, Newton, NC). After treatment, samples were stored at −80°C. All samples were subject to a single freeze–thaw cycle to break down mucopolysaccharides that can interfere with pipetting (Worthman et al., 1990). On the day of assay, samples were centrifuged (1500 g, 15 min) to remove particulate matter, then clear samples were transferred into appropriate test tubes or wells. To minimize error, all samples were tested in duplicate; duplicates that varied by more than 5% error were repeat tested, and the average of the duplicates for each sample was used in the analyses. All samples from each individual subject (cotton-treated and untreated samples) were tested in the same assay run.

2.2. Immunoassay protocols

Clear and cotton-treated saliva samples were assayed (without separation or extraction) for DHEA–S, DHEA, cortisol, cotinine, testosterone, estradiol, progesterone and sIgA. Performance characteristics for each assay are presented in Table 1. The manufacturer’s recommended EIA protocols for use with saliva were followed without modification for DHEA, cortisol, testosterone, and progesterone (Salimetrics, State College, PA), sIgA (American Laboratory Products, Windham, NH), and DHEA–S (ELISA) (Diagnostic Systems Laboratories, Webster, TX). Serum RIA protocols were modified for use with saliva for DHEA (see Granger et al., 1999b), testosterone (see Granger et al., 1999a), and estradiol (Shirtcliff et al., 2000). Salivary cotinine was assayed using a modification of the Diagnostic Products Corporation (DPC, Los Angeles CA) “Coat-a-count RIA kit” for urinary cotinine (see Granger and Schwartz, 1998). All assay protocols are available from the kit manufacturers or from the authors on request.
Table 1
Assay sensitivity and reliability

<table>
<thead>
<tr>
<th></th>
<th>Analytical sensitivity</th>
<th>Intraassay CV%</th>
<th>Inter assay CV%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Radioimmunoassays</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotinine</td>
<td>1 ng/ml</td>
<td>7.83</td>
<td>11.05</td>
</tr>
<tr>
<td>DHEA</td>
<td>4 pg/ml</td>
<td>4.05</td>
<td>12.57</td>
</tr>
<tr>
<td>DHEA-S</td>
<td>0.1 ng/ml</td>
<td>9.00&lt;sup&gt;a&lt;/sup&gt;</td>
<td>13.00&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Estradiol</td>
<td>0.25 pg/ml</td>
<td>6.45</td>
<td>9.00</td>
</tr>
<tr>
<td>Testosterone</td>
<td>0.8 pg/ml</td>
<td>5.00</td>
<td>7.00</td>
</tr>
<tr>
<td><strong>Enzymeimmunoassays</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortisol</td>
<td>0.007 µg/dl</td>
<td>4.13</td>
<td>8.89</td>
</tr>
<tr>
<td>DHEA</td>
<td>10 pg/ml</td>
<td>3.45</td>
<td>4.90</td>
</tr>
<tr>
<td>Testosterone</td>
<td>1.5 pg/ml</td>
<td>3.30</td>
<td>6.70</td>
</tr>
<tr>
<td>Progesterone</td>
<td>10 pg/ml</td>
<td>8.35</td>
<td>9.40</td>
</tr>
<tr>
<td>sIgA</td>
<td>20 µg/ml</td>
<td>2.79</td>
<td>11.61</td>
</tr>
</tbody>
</table>

<sup>a</sup> Values are conservative estimations based on the largest coefficient of variation.

2.3. Analytical strategy

A mixed-model ANOVA design was used to assess whether cotton absorbent materials had systematic effects across, or specific effects by marker type. Paired t-tests (with and without correction for alpha-wise error) compared cotton vs. no-cotton methods for each marker separately. Single sample t-tests were used to determine whether the magnitude of change in saliva levels attributable to use of cotton collection methods was significantly greater than each assay’s inter-assay coefficient of variation (i.e., random error expected in results due to subtle differences in assay performance across runs). Rank and linear correlations were computed to assess whether the order of individual scores within each marker’s distribution was conserved when samples were collected using the cotton vs. no-cotton methods. Finally, regression equations were used to determine whether saliva results generated using cotton or no-cotton methods were better estimates of blood levels.

3. Results

3.1. Quantitative estimates of the levels of salivary biomarkers

The first analysis used a five (marker type: DHEA, cortisol, testosterone, progesterone, sIgA) by two (collection method: Cotton vs. No-cotton) mixed-model ANOVA for the EIA protocols. The second analysis used a five (marker type: DHEA, DHEA–S, estradiol, cotinine, and testosterone) by two (collection method) mixed-model ANOVA for the RIA protocols. In both ANOVAs marker type was included as a blocking factor, and collection method was the repeated measure. The main effect of collection method is significant for EIA, $F(1,25)=174.41$, $P<0.0005$, and RIA protocols, $F(1,45)=650.49$, $P<0.0005$. The marker type by collection method interac-
tions are also highly significant for both EIA, $F(2,25)=76.30, P<0.0005$, and RIA, $F(4,45)=543.11, P<0.0005$, protocols. The source of the interactions was decomposed by paired $t$-tests (two-tailed) comparing no-cotton to cotton methods for each marker separately. Prior to correction for alpha-wise error, there are statistically significant effects of collection method ($P<0.05$) on assay results for all the salivary markers except cotinine. After Bonferroni correction ($P<0.005$), significant differences by collection method are observed for DHEA, estradiol, testosterone, progesterone, and sIgA, but not for cortisol, DHEA–S, and cotinine. The magnitude and direction of the effect is highly specific to marker type. See Table 2 for means and standard deviations.

Single sample $t$-tests (with correction for alpha-wise error, $P<0.005$) were used to determine whether the magnitude of change in saliva levels attributable to the use of cotton collection methods was significantly greater than each assay’s inter-assay coefficient of variation. The impact of collection method on assay results was significantly greater ($Ps<0.005$) than the unsystematic error in assay performance in every case, with cotinine (RIA), cortisol (EIA), sIgA (EIA), DHEA–S (RIA) and testosterone (EIA) being the exceptions.

### 3.2. Individual scores within each salivary biomarker’s distribution

Rank-order and linear correlations were computed between cotton and no-cotton conditions separately for each marker to evaluate the degree to which individual

### Table 2

<table>
<thead>
<tr>
<th>Assay type</th>
<th>Marker</th>
<th>$n$</th>
<th>No-cotton</th>
<th>Cotton</th>
<th>Rank Order $R$</th>
<th>Linear $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radioimmunoassay</td>
<td>Cotinine (ng/ml)</td>
<td>13</td>
<td>56.21 (76.30)</td>
<td>63.98 (97.97)</td>
<td>0.95**</td>
<td>0.99**</td>
</tr>
<tr>
<td></td>
<td>DHEA (pg/ml)</td>
<td>6</td>
<td>94.02 (20.14)</td>
<td>1211.61** (64.10)</td>
<td>0.26</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>DHEA-S (ng/ml)</td>
<td>5</td>
<td>1.57 (0.33)</td>
<td>1.06 (0.23)</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Estradiol (pg/ml)</td>
<td>5</td>
<td>0.74 (0.45)</td>
<td>3.30** (0.27)</td>
<td>−0.30</td>
<td>−0.30</td>
</tr>
<tr>
<td></td>
<td>Testosterone (pg/ml)</td>
<td>21</td>
<td>18.71 (5.01)</td>
<td>40.44** (9.84)</td>
<td>0.95**</td>
<td>0.90**</td>
</tr>
<tr>
<td>Enzymeimmunoassay</td>
<td>Cortisol (ug/dl)</td>
<td>11</td>
<td>0.15 (0.03)</td>
<td>0.12 (0.02)</td>
<td>0.93**</td>
<td>0.96**</td>
</tr>
<tr>
<td></td>
<td>DHEA (pg/ml)</td>
<td>4</td>
<td>40.43 (10.16)</td>
<td>571.70** (133.20)</td>
<td>1.0**</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Testosterone (pg/ml)</td>
<td>4</td>
<td>79.10 (57.82)</td>
<td>229.60** (40.06)</td>
<td>0.60</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>Progesterone (pg/ml)</td>
<td>5</td>
<td>182.65 (230.27)</td>
<td>&gt;1000** (NA)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>sIgA (µg/ml)</td>
<td>6</td>
<td>176.45 (62.33)</td>
<td>48.95** (51.67)</td>
<td>0.26</td>
<td>0.43</td>
</tr>
</tbody>
</table>

* $** P<0.005$, paired $t$-tests (two-tailed) between control and cotton methods, Rank Order Correlations, and Linear Correlations. NA — All results for salivary progesterone in the cotton condition exceeded the value of the highest standard in the assay.
differences in assay results were conserved across the collection methods. As can be seen in Table 2, in addition to changes in average levels of markers (as described above), unsystematic error is introduced in the individual scores for some markers more than others. That is, the order of the individual scores for cortisol (EIA), cotinine (RIA), testosterone (RIA), and DHEA (EIA) is conserved regardless of the collection method. In contrast, use of cotton materials during sample collection introduces unsystematic error among the individual scores in the distributions for DHEA (RIA), DHEA–S (RIA), estradiol (RIA), Testosterone (EIA), and sIgA (EIA).

3.3. Predicting blood levels using salivary levels

To determine whether saliva collected without cotton yielded better estimates of blood levels of DHEA than cotton methods, we collected matched serum and saliva from 28 adults (14 males). When saliva is collected without cotton absorbent materials, the serum–saliva correlation for DHEA is highly significant, $r(26)=0.859$, $P<0.0005$, but the correlation becomes non-significant when saliva is collected using the cotton method, $r(26)=0.209$, $P=0.285$, (see Fig. 1). Two regression equations were computed predicting serum DHEA levels. In the first equation, saliva DHEA levels from samples collected by the no-cotton method where entered first, followed by the saliva levels from samples collected using the cotton-method. Results revealed that the no-cotton method saliva levels explained a significant percent of the variance in serum levels ($r^2(26)=0.727$, $P<0.0005$), and the cotton-method saliva levels did not contribute uniquely to the prediction. In the second equation, the cotton-method saliva levels were entered first, followed by the no-cotton method saliva levels. Results confirm that the cotton-method levels do not contribute significantly to the prediction of serum levels ($r^2(25)=0.044$), and the no-cotton method levels contribute uniquely to the prediction above and beyond the variance accounted for by the cotton-method levels (change in $r^2=0.713$, $P<0.0005$). Saliva assay results generated...
from samples collected without cotton explain significantly more of the variance in the concentration of DHEA in the general circulation than do salivary results generated from samples collected using cotton-absorbent materials.

4. Discussion

The use of cotton-based absorbent materials has the potential to profoundly affect the results of immunoassays for certain salivary biomarkers. Immunoassay results for salivary testosterone, DHEA, progesterone, and estradiol are higher, whereas sIgA results are significantly lower, when samples are collected using cotton absorbent materials compared to samples collected without cotton. The magnitude of the error introduced is significantly greater than would be expected by variation from run-to-run in assay performance. The method also affects unsystematic error such that assumptions regarding saliva–serum correspondence may not be valid. In contrast, results for markers such as cortisol, DHEA–S, and cotinine are not affected by the cotton collection procedure. The effects of the cotton material are similar regardless of whether the assay used is a radio- or enzyme-immunoassay.

The differential effect by marker type suggests possible mechanisms. Cotton-related elevations in assay results might be due to the presence of an interfering substance in the cotton. Previous studies suggest that plant hormones might cross-react with antibodies used for certain assays (Dabbs, 1991; Granger et al., 1999a). That the effect is more pronounced for some assays than others suggests that the interference may be due to non-specific binding or cross-linking rather than cross-reactivity with the specific antibody for the marker in question. Alternatively, the cotton may filter out an interfering substance. Cotton-related reductions in assay results raise the possibility that some of the molecules in the sample are adhering to the cotton fibers. In subsequent analyses, we have confirmed that neither the magnitude nor direction of the effect of cotton is associated with the molecular weight of the markers included in this study. Identification of the source of the explanation is beyond the purpose of this report but would seem worthwhile in future studies aimed at designing collection alternatives.

This pattern of effects warns investigators that, depending on their markers of interest, cotton may potentially interfere with immunoassay results. Appropriate collection procedures need to be taken into consideration for different types of markers. An important take home message is that it may not be feasible to assume that saliva samples collected using one particular method can be assayed for multiple markers. Researchers with intentions of analyzing multiple biomarkers in saliva should conduct pilot testing to verify that different collection procedures do not interfere with each anticipated assay procedure.

With respect to previously reported studies that may have collected samples without attention to these details, the news is not good. Our findings suggest that the measures of salivary markers reported in such studies may be compromised, and the magnitude of the marker-behavior associations reported may be underestimated. Less obvious, but equally important are implications for investigators who are archiving
samples to be assayed for a range of markers in the future. If the range of assays available for salivary use continues to increase, sample archives constructed using restrictive collection techniques seriously compromises the utility of the archive.

Although the present study yielded some potentially important findings, it provides only a partial picture, and several limitations qualify the interpretation of the data. It is important to note that the assays employed did not employ separation or extraction steps. Thus, the generalization of these findings to assay protocols that use such procedures would be highly speculative. Investigators interested in specific markers might also be well advised to re-do these comparisons in their own labs using their preferred protocols, perhaps employing larger sample sizes.

In conclusion, the immunoassay of saliva collected using cotton-based materials to absorb sample has the potential to be a source of systematic and unsystematic error in measurement. The success of the next generation of studies will depend on variance in salivary biomarker levels that reflect true individual differences rather than the effects of extraneous factors.

Acknowledgements

This study was supported by The Pennsylvania State University Behavioral Endocrinology Laboratory. Technical support was provided by Jodi Heaton and Thema Nicholson. We gratefully acknowledge the contribution of reagents by Rusty Nicar and Prasanna Harihar of Diagnostic Systems Laboratories (Webster, TX).

References