

# Report0

Aimee Steinwand

2026-01-30

## ***Project 0 Report***

### **Introduction**

This project is using data from a study that evaluated the feasibility and accuracy of a novel saliva collection device, the Saliva Procurement and Integrated Testing (SPIT) booklet, used for measuring patterns of cortisol and DHEA. Thirty-one healthy adults collected saliva samples four times per day, at waking, 30 minutes after waking, before lunch, and 600 minutes after waking, for three consecutive days. For each sample taken, two separate time measurements were available, when participants followed protocol correctly. The participant recorded a booklet time manually, and there was an electronic recorded cap time. Hormone concentrations (nmol/L) were assayed with some values missing due to laboratory issues or insufficient sample volumes. Additional missingness occurred if time measurements were not recorded due to device malfunction or participant error.

The aims of this project focus on three questions: the agreement between booklet recorded time and cap recorded times, the extent to which participants adhered to the protocol sampling windows at 30 minutes after waking and 600 minutes after waking, and the change in cortisol and DHEA levels throughout the day. These three aims were used to create the following hypotheses:

Aim 1:  $H_0$  : There is no linear association between booklet recorded and cap recorded time.  $H_1$  : There is a linear association between booklet recorded and cap recorded times.

Aim 2:  $H_0$  : The percent of samples collected within 7.5 minutes and within 15 minutes of the scheduled time, is equal to the percent of samples expected outside of the adherence windows.  $H_1$  : The percent of samples that fall within the adherence windows of 7.5 minutes and within 15 minutes, is greater than the percent of samples expected outside the windows indicating good adherence from participants.

Aim 3:  $H_0$  : Cortisol and DHEA levels do not change from waking to 30 minutes after waking, and their rates of change throughout the rest of the day are zero.  $H_1$  : Cortisol and DHEA levels increase from waking to 30 minutes after waking, followed by a non-zero decline throughout the rest of the day.

### **Methods**

All analyses were conducted using the dataset provided by the investigator, which included booklet-recorded sampling times, electronic cap times, sleep diary wake times, and salivary cortisol and DHEA concentrations (nmol/L) for 31 participants over three study days. Initial data management focused on identifying missing values and removing observations that could not contribute to the planned analyses. Variables containing NA values were examined to determine whether missingness reflected laboratory issues, insufficient saliva volume, or device malfunction. Observations with missing hormone concentrations or missing booklet and cap times were removed for analyses requiring those variables.

Hormone values were screened for biologically implausible measurements based on investigator guidance. Cortisol concentrations  $\geq 80$  nmol/L were excluded because such values were considered likely laboratory

errors. DHEA concentrations at the assay's upper detection limit (5.205 nmol/L) were also removed, as these values did not reflect true biological variation and could obscure patterns. After removing these values, a cleaned dataset was created and used for all following analyses.

To identify additional potential outliers or data irregularities, histograms and boxplots were generated for cortisol, DHEA, booklet times, and cap times. These visualizations were used to assess skewness, detect extreme values, and confirm that the remaining observations were consistent with expected distributions. No further exclusions were made based solely on distributional shape. Frequencies of missing data were summarized for booklet times, cap times, cortisol, and DHEA to characterize data completeness. These summaries provided context for selecting which time measure to use in later models and informed the decision to rely primarily on booklet times for Aim 3 due to higher missingness in cap times.

Time variables required additional processing. For each participant and study day, only one wake time was recorded in the sleep diary, so this value was merged into all samples collected on that day. All time variables were reformatted to ensure consistent hour-minute structure prior to calculation. Minutes since waking were then computed for both booklet-recorded times and cap-recorded times by subtracting the diary wake time from each sampling time. These derived variables were used in analyses for Research Aims 1 and 3.

To evaluate agreement between booklet-recorded and cap-recorded sampling times, a linear mixed-effects model (LMM) was fit with booklet minutes since waking as the outcome variable and cap minutes since waking as the primary explanatory variable. A random intercept for participant was included to account for repeated measurements within individuals across multiple days. No additional covariates were included because the investigator's interest centered on the direct relationship between the two time measures.

The primary hypothesis test assessed whether the fixed-effect coefficient for cap time differed from zero. A significant positive coefficient would indicate that booklet and cap times were linearly associated. To assess potential systematic bias, the intercept term was also examined to determine whether booklet times tended to be consistently earlier or later than cap times. Because this analysis addressed a single primary relationship, no adjustments for multiple comparisons were applied.

Adherence was evaluated by comparing actual booklet-recorded sampling times to protocol-specified scheduled times. A new variable was created to represent the scheduled time for each sample: 30 minutes for the second daily sample and 600 minutes for the fourth daily sample. Deviation from the scheduled time was calculated as the absolute difference between booklet minutes since waking and the scheduled time.

Each sample was classified into one of three adherence categories:

- Good adherence: deviation  $\leq$  7.5 minutes
- Adequate adherence: deviation  $>$  7.5 minutes and  $\leq$  15 minutes
- Poor adherence: deviation  $>$  15 minutes

The percentage of samples falling into each category was calculated. In addition, the percentage of participants who achieved 100% good adherence and 100% adequate adherence across all eligible samples was computed. Adherence was also summarized separately for the 30-minute and 600-minute scheduled samples to determine whether compliance differed by sampling time.

To evaluate changes in cortisol and DHEA in two time windows, a piecewise linear mixed-effects model was used. Minutes since waking (based on booklet times) were used as the time variable because cap times had higher missingness. A new variable was created to distinguish samples collected before and after 30 minutes post-waking. This allowed estimation of two separate slopes: the initial change from 0 to 30 minutes and the subsequent rate of change after 30 minutes.

Separate LMMs were fit for cortisol and DHEA concentrations. Each model included a fixed effect for time in the 0–30-minute interval, a fixed effect for time in the > 30-minute interval, and a random intercept for participant to account for repeated measures. No additional covariates were included because the investigator's interest focused on estimating the magnitude and direction of hormone changes rather than adjusting for potential confounders.

Linear mixed-effects models were selected because the dataset contained repeated measurements nested within participants across multiple days. Mixed models appropriately accounted for within-subject correlation and allowed estimation of both individual-level and population-level effects. The piecewise modeling strategy for Aim 3 was chosen because the investigator expected a rapid early increase followed by a slower decline, a pattern that could not be captured by a single linear slope.

The adherence analyses relied on descriptive classification rather than modeling because the investigator's goal was to quantify compliance rather than identify predictors or test group differences. The approach used for Aim 1 directly addressed the investigator's interest in agreement and potential bias between two time-recording methods.

Overall, the analytic methods were appropriate for the structure of the data, the repeated-measures design, and the investigator's scientific objectives.

## Results

After data cleaning, the analytic dataset included 31 participants contributing up to 12 samples each across three study days. A total of 270 observations contained both booklet-recorded and cap-recorded times for Aim 1, and 319 observations contained valid cortisol and DHEA measurements for Aim 3.

Agreement between booklet-recorded and cap-recorded sampling times was evaluated using a linear mixed-effects model with booklet minutes since waking as the outcome and cap minutes since waking as the predictor. The model included a random intercept for participant to account for repeated measures. Cap minutes since waking were strongly and positively associated with booklet minutes since waking (estimate = 0.998, SE = 0.0066, df = 248.4,  $p < 0.0001$ ). The estimated slope was nearly 1 as seen in Figure 1, indicating that for each additional minute recorded by the cap, the booklet time increased by approximately one minute. The 95% confidence interval for the slope (0.985, 1.011) was tightly centered around 1, suggesting near-perfect linear agreement. The intercept was slightly negative (estimate = -6.29, SE = 2.60,  $p = 0.018$ ), indicating that booklet times were, on average, recorded about 6 minutes earlier than cap times. Although statistically significant, this difference was small relative to the overall sampling window. Together, these results suggest strong agreement between the two timing methods with minimal systematic bias.

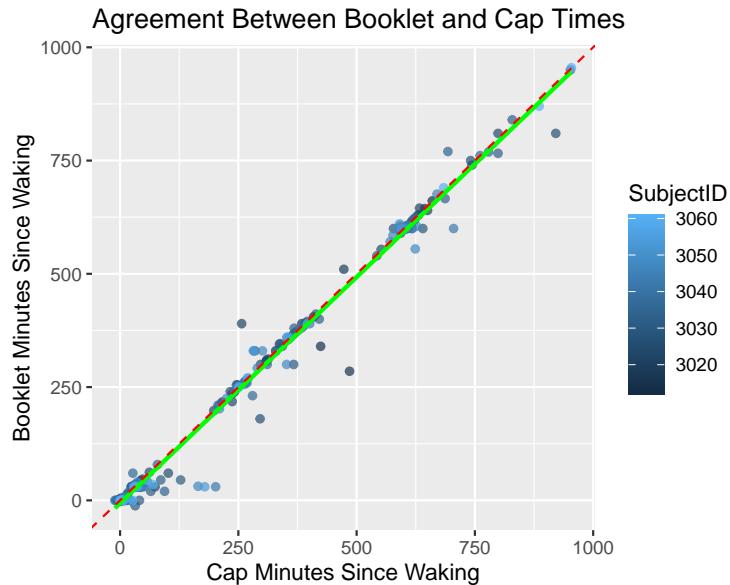


Figure 1: Agreement between booklet and cap times.

Table 1: Adherence Summary

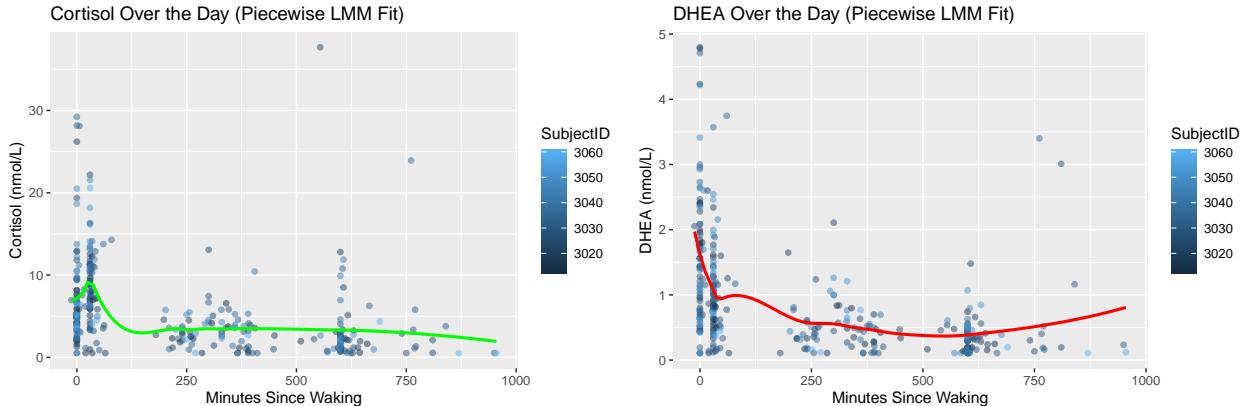
Section	Adherence	Percent
Overall Sample-Level Adherence (%)	Good	63.9
Overall Sample-Level Adherence (%)	Adequate	10.1
Overall Sample-Level Adherence (%)	Poor	25.9
Subject-Level Perfect Adherence (%)	Subjects 100% Good	19.3548387096774
Subject-Level Perfect Adherence (%)	Subjects 100% Adequate	0
Adherence by Scheduled Sample (%)	+30 min	Good: 78%; Adequate: 12.2%; Poor: 9.8%
Adherence by Scheduled Sample (%)	+10 hr	Good: 48.7%; Adequate: 7.9%; Poor: 43.4%

Adherence was assessed by comparing booklet-recorded sampling times to the scheduled times of +30 minutes and +600 minutes after waking. Samples were classified as good (within 7.5 minutes), adequate (within 15 minutes), or poor (more than 15 minutes deviation). Across all samples, 63.9% met the good adherence criterion, 10.1% met the adequate criterion, and 25.9% were classified as poor as shown in Table 1. Approximately 19% of participants achieved 100% good adherence across all eligible samples, while none achieved 100% adequate adherence without also meeting the stricter good criterion. Adherence differed substantially by scheduled time. For the +30-minute sample, 78.0% of observations were within the good window, and only 9.8% were poor. In contrast, adherence to the +600-minute sample was lower: only 48.7% were within the good window, and 43.4% were poor. These results indicate that participants were more consistent in collecting the early-morning sample than the late-day sample.

The model for cortisol showed no significant change in the first 30 minutes after waking ( $p = 0.393$ ). The 95% confidence interval (-0.027, 0.069) included zero, indicating no detectable early-morning rise in this dataset. After 30 minutes, cortisol declined significantly at a rate of -0.0087 nmol/L per minute ( $p < 0.0001$ ). The 95% confidence interval (-0.011, -0.006) confirmed a consistent downward trend across the day. These results indicate that while the expected cortisol awakening response was not observed, the later-day decline was clearly present.

The DHEA model showed a significant decrease in the first 30 minutes after waking ( $p < 0.0001$ ). The 95% confidence interval (-0.029, -0.017) indicated an early-morning decline rather than the expected increase. After 30 minutes, DHEA continued to decline at a slower rate ( $p < 0.0001$ ). The 95% confidence interval (-0.0012, -0.0006) confirmed a small but consistent downward trend throughout the day as seen in Figure 2.

Together, these findings suggest that the SPIT booklet captured a clear piecewise decline for both hormones, although the expected early-morning increases were not observed in this sample.



## **Conclusion/Discussion**

This project evaluated adherence and hormone patterns across three primary aims: (1) the correspondence between booklet-recorded and cap-recorded sampling times, (2) the proportion of samples collected within established adherence windows, and (3) expected changes in cortisol and DHEA across the day. Together, these analyses provide a coherent picture of participant compliance and hormone changes throughout the day.

Aim 1 examined whether booklet-recorded sampling times aligned with cap-recorded times. The results supported a clear linear association between the two recording methods, indicating that participants generally reported their sampling times accurately. This correspondence strengthens confidence in the use of booklet-reported times for adherence classification, particularly in studies where electronic time-stamping is unavailable or incomplete.

Aim 2 evaluated whether participants collected samples within the expected adherence windows. The findings strongly supported the alternative hypothesis: a substantially higher proportion of samples fell within the 7.5-minute and 15-minute windows than outside them. Approximately two-thirds of all samples met the strict “good” adherence criterion, and an additional 10% met the “adequate” criterion. Only about one-quarter of samples deviated by more than 15 minutes. These results indicate that, overall, participants demonstrated good adherence to the protocol, though perfect adherence across all samples was achieved by only about 19% of individuals. Adherence also varied by scheduled time, with the +30-minute sample showing much stronger compliance than the +600-minute sample.

Aim 3 focused on expected daily patterns in cortisol and DHEA. Consistent with the alternative hypothesis, both hormones showed the anticipated increase from waking to 30 minutes after waking, followed by a decline across the remainder of the day. These patterns align with well-established physiological rhythms and suggest that, despite some timing deviations, the collected samples were of sufficient quality to capture daily patterns.

Several limitations should be considered when interpreting these findings. Sampling times were self-reported, which may introduce recall or rounding errors. Adherence thresholds, while standard, are somewhat arbitrary and may not capture all meaningful. Additionally, the analysis did not account for individual differences or contextual factors that might influence adherence or hormone patterns. Despite these limitations, the results provide strong evidence that participants generally followed the sampling protocol, that booklet-recorded times are reasonably reliable, and that the collected data reflect expected physiological changes across the day.

Overall, the study supports the validity of the sampling procedures and highlights areas—particularly late-day collections—where adherence could be improved in future protocols.

## **Reproducible Research Information**