1 Introduction

This analysis evaluates a novel saliva collection device—the Saliva Procurement and Integrated Testing (SPIT) booklet—designed for at-home collection of salivary cortisol and DHEA on filter paper strips. The dataset includes 31 healthy subjects who collected saliva samples four times daily (at waking, 30 minutes post-waking, before lunch, and 600 minutes post-waking) over three consecutive days. The SPIT booklet was stored in a bottle with an electronic monitoring (MEMs) cap that independently recorded sampling times. For each sample, the dataset contains subject-recorded times (Booklet), MEMs-recorded times, sleep diary wake times, and hormone concentrations in nmol/L units.

The investigator posed three research questions. First, do subject-recorded sampling times agree with electronically-recorded times, and is there systematic bias between the two methods? Second, do subjects adhere to the protocol-specified timing windows for the +30 minute and +600 minute samples? Third, do cortisol and DHEA exhibit the expected diurnal patterns—an initial increase from waking to 30 minutes post-waking followed by a gradual decline?

These questions translate into the following statistical hypotheses. For Question 1, the null hypothesis is that Booklet-recorded times since waking show no systematic bias relative to MEMs-recorded times; the alternative is that systematic bias exists. For Question 2, adherence is assessed descriptively by calculating the proportion of samples collected within ± 7.5 minutes and ± 15 minutes of protocol-specified times. For Question 3, the null hypotheses are: (a) there is no change in hormone levels between waking and 30 minutes post-waking, and (b) there is no decline in hormone levels from 30 minutes post-waking through 600 minutes post-waking; the alternatives posit significant changes during each period. All hypothesis tests use $\alpha = 0.05$.

2 Methods

2.1 Data Cleaning

For cortisol and DHEA, only measurements in nmol/L units were used in analysis. In analysis to address Q3, cortisol values exceeding 80 nmol/L were excluded as likely lab errors. For DHEA, values at the upper detection limit (5.205 nmol/L) were flagged; subjects with multiple observations at this limit were identified for potential exclusion from the hormone analysis due to possible underlying health conditions.

Clock times recorded by subjects (Booklet), the electronic cap (MEMs), and sleep diary wake times were converted from hh:mm format to minutes since midnight. Wake time, recorded only at each day's first sample, was carried forward to all samples within each subject-day. Two derived variables were calculated: Booklet time since waking and MEMs time since waking (clock time minus wake time, in minutes).

2.2 Statistical Analysis

2.2.1 Question 1: Agreement Between Booklet and MEMs Times

Agreement between subject-recorded (Booklet) and electronically-recorded (MEMs) times since waking was assessed using a linear mixed effects model. The outcome variable was Booklet time since waking (minutes), with MEMs time since waking as the primary explanatory variable. A random intercept for subject was included to account for repeated measures within individuals. Perfect agreement would yield a slope of 1 and intercept of 0 and systematic bias is accounted by the fixed intercept.

2.2.2 Question 2: Adherence to Protocol Timing

Adherence was evaluated for the two protocol-specified sampling times: Sample 2 (target: 30 minutes post-waking) and Sample 4 (target: 600 minutes post-waking). Using Booklet times since waking as the objective measure, the deviation from target was calculated for each observation. Adherence was quantified as the percentage of samples collected within ± 7.5 minutes and within ± 15 minutes of the target time. Descriptive statistics (mean deviation, standard deviation, median, range) were computed for each sample type. A scatterplot showing the relationship is in the Appendix section.

2.2.3 Question 3: Changes in Cortisol and DHEA Over Time

Diurnal patterns in cortisol and DHEA (nmol/L) were modeled using linear mixed effects models with MEMs-recorded time since waking as the primary explanatory variable. A random intercept for subject was included to account for within-subject correlation from repeated measures.

To capture the expected biphasic pattern—an initial rise from waking to 30 minutes (cortisol awakening response) followed by a gradual decline—piecewise linear models were fit with a knot at 30 minutes. The model included two time components: time from 0 to 30 minutes (slope representing the awakening response) and time beyond 30 minutes (slope representing the subsequent decline). Model diagnostics included residual plots, Q-Q plots for normality assessment.

For cortisol, the primary hypotheses tested were: (1) whether the slope from 0 to 30 minutes significantly differed from zero (awakening response), and (2) whether the slope after 30 minutes significantly differed from zero (diurnal decline). The same hypotheses were tested for DHEA. Point estimates with 95% confidence intervals are reported regardless of statistical significance, per the investigator's request to compare effect sizes with published literature.

3 Results

3.1 Descriptive Statistics and Missing Data

Table 1 summarizes study sample characteristics. The dataset included 31 subjects contributing 372 total observations over 3 days of collection. Mean Booklet time since waking was 239.4 minutes (SD = 263.5) and mean MEMs time since waking was 272.0 minutes (SD = 272.3). Mean cortisol was 5.95 nmol/L (SD = 6.95) and mean DHEA was 0.98 nmol/L (SD = 1.03).

MEMs clock times had the highest missingness (16.4%), likely due to cap malfunction or subjects failing to return the booklet to the bottle between samples. Booklet clock times were missing for 9.4% of observations, while wake time was complete. Cortisol and DHEA were each missing for 1.3% of observations. For the Q3 hormone analysis, one cortisol value (0.3%) was excluded as a lab error and six DHEA values (1.6%) were excluded for being at the detection limit. One subject (ID 3037) was excluded from the DHEA analysis due to multiple values at the detection limit.

Table 1: Study Sample Characteristics

Characteristic	Value
Sample Size	
Subjects, n	31
Total observations, n	372
Days of collection per subject	3
Descriptive Statistics, mean (SD)	
Booklet time since waking (min)	239.4 (263.5)
MEMs time since waking (min)	272.0 (272.3)
Cortisol (nmol/L)	5.95 (6.95)
DHEA (nmol/L)	0.98 (1.03)
Missing Data, n (%)	
Booklet clock time	35 (9.4%)
MEMs clock time	61 (16.4%)
Wake time	0 (0.0%)
Cortisol	5 (1.3%)
DHEA	5 (1.3%)
Data Exclusions for Q3 Analysis, n (%)	
Cortisol $>$ 80 nmol/L (lab error)	1 (0.3%)
DHEA at detection limit (5.205 nmol/L)	6 (1.6%)

Note: Time since waking calculated from sleep diary wake time and recorded clock time. One subject (ID 3037) was excluded from Q3 DHEA analysis due to multiple values at detection limit.

3.2 Question 1: Agreement Between Booklet and MEMs Times

The linear mixed model assessing agreement between Booklet and MEMs recording times demonstrated strong agreement between the two methods (Table 2). The slope for MEMs time was 0.995 (SE = 0.007, 95% CI: 0.981 to 1.009, $p < 0.001$), indicating a near-perfect 1:1 relationship—for each additional minute recorded by the electronic cap, subjects recorded approximately one additional minute in the booklet.

However, the intercept of -6.502 minutes (SE = 2.913, 95% CI: -12.21 to -0.79, p = 0.026) was significantly different from zero, indicating systematic bias. The negative intercept suggests that subjects tended to record sampling times approximately 6.5 minutes earlier than the times captured by the electronic monitoring cap.

The random effects variance components (Table 3) showed that most variability in the Booklet-MEMs relationship occurred within subjects (residual SD = 31.46 minutes) rather than between subjects (subject SD = 6.64 minutes). The intraclass correlation coefficient (ICC) of 0.043 indicates that only 4.3% of the total variance in Booklet times (after accounting for MEMs times) was attributable to between-subject differences.

In summary, subject-recorded and electronically-recorded times since waking were strongly associated, but subjects systematically recorded times approximately 6.5 minutes earlier than the electronic cap. This bias, while statistically significant, is relatively modest in magnitude given the range of sampling times (0 to 600+ minutes).

Table 2: Summary of Agreement Analysis Between Booklet and MEMs Recording Times

Metric	Value	Interpretation
Sample Size		
Number of observations	285	
Number of subjects	31	
Proportional Bias (Slope)		
Slope estimate	0.9949	Close to 1 indicates proportional agreement
Slope 95% CI	(0.9811, 1.0087)	
Test: Slope $= 1$ (t-value)	-0.72	
Test: Slope $= 1$ (p-value)	0.471	Not significant: no proportional bias
Constant Bias (Intercept)		
Intercept estimate	-6.50 min	Booklet times recorded earlier than MEMs
Intercept 95% CI	$(-12.21, -0.79)$	
Variance Components		
Between-subject variance	44.08	Between-subject variability in bias
Residual variance	989.60	Within-subject variability
ICC	0.043 (4.3%)	Most variation is within-subject

Table 3: Adherence to Protocol-Specified Sampling Times

	+30 min Sample		+10 hours Sample	
Window	n/N	% (95% CI)	n/N	% (95% CI)
$\pm$7.5 min	66/87	75.9 (65.5, 84.4)	36/80	45.0 (33.8, 56.5)
$\pm$15 min	76/87	87.4 (78.5, 93.5)	44/80	55.0 (43.5, 66.2)

Note: Adherence calculated using subject-recorded Booklet time since waking.

3.3 Question 2: Adherence to Protocol Timing

Table 3 summarizes adherence to protocol-specified sampling times using subject-recorded Booklet times since waking. For the +30 minute sample, 75.9% (95% CI: 65.5%, 84.4%) of observations were collected within $\pm$7.5 minutes of the target, and 87.4% (95% CI: 78.5%, 93.5%) were collected within $\pm$15 minutes. Adherence was lower for the +10 hour sample: 45.0% (95% CI: 33.8%, 56.5%) of observations fell within $\pm$7.5 minutes of the target, and 55.0% (95% CI: 43.5%, 66.2%) were within $\pm$15 minutes.

Overall, subjects adhered more closely to the +30 minute sampling window than the +10 hour window. The majority of +30 minute samples (87.4%) met the $\pm$15 minute tolerance, whereas only about half (55.0%) of +10 hour samples met this criterion, suggesting that maintaining precise timing for samples later in the day is more challenging in at-home collection settings.

3.4 Question 3: Changes in Cortisol and DHEA Over Time

Table 4 presents results from the piecewise linear mixed models examining diurnal patterns in log-transformed cortisol and DHEA. Log transformation was applied to address right-skewed distributions typical of hormone data.

3.4.1 Cortisol

The estimated geometric mean cortisol at waking was $e^{1.719} = 5.58$ nmol/L ($p < 0.001$). For the first question, the estimated slope from waking to 30 minutes was 0.0060 log(nmol/L) per minute. Back-transforming, this corresponds to a multiplicative factor of $e^{0.0060} = 1.006$ per minute, meaning cortisol increases by approximately 0.6% per minute; however, this was not statistically significant (p = 0.241). For the second question, cortisol declined significantly after 30 minutes with a slope of -0.0022 log(nmol/L) per minute ($p < 0.001$). Back-transforming, this corresponds to a multiplicative factor of $e^{-0.0022} = 0.998$ per minute, meaning cortisol decreases by approximately 0.22% per minute.

The ICC of 14.1% indicates that a modest proportion of cortisol variability was attributable to between-subject differences.

3.4.2 DHEA

The estimated geometric mean DHEA at waking was $e^{0.451} = 1.57$ nmol/L (p = 0.002). For the first question, DHEA did not show an increase from waking to 30 minutes; instead, it significantly decreased with a slope of -0.0258 log(nmol/L) per minute ($p < 0.001$). Back-transforming, this corresponds to a multiplicative factor of $e^{-0.0258} = 0.975$ per minute, meaning DHEA decreases by approximately 2.5% per minute. For the second question, DHEA continued to decline after 30 minutes with a slope of -0.0015 log(nmol/L) per minute ($p < 0.001$). Back-transforming, this corresponds to a multiplicative factor of $e^{-0.0015} = 0.9985$ per minute, meaning DHEA decreases by approximately 0.15% per minute.

The ICC of 44.2% indicates that a substantial proportion of DHEA variability was attributable to between-subject differences, suggesting more stable individual differences in DHEA levels compared to cortisol.

Table 4: Piecewise Linear Mixed Model Results for Log-Transformed Cortisol and DHEA

Parameter	log(Cortisol)	log(DHEA)
Fixed Effects		
Intercept (at waking)	1.719 ($p < 0.001$)	0.451 (p = 0.002)
Time1 (0–30 min slope)	0.0060 (p = 0.241)	-0.0258 ($p < 0.001$)
Time2 (post-30 min slope)	-0.0022 ($p < 0.001$)	-0.0015 ($p < 0.001$)
Random Effects		
Subject variance	0.090	0.286
Residual variance	0.547	0.362
ICC	14.1%	44.2%

Note: Model: log(Hormone) $\sim$ Time1 + Time2 + (1|SubjectID). Time1 = min(Time, 30); Time2 = max(Time $-$ 30, 0). Knot at 30 minutes after waking. N = 282 observations from 30 subjects. Slopes are in log(nmol/L) per minute. ICC = Intraclass correlation coefficient.

3.4.3 Summary

The SPIT booklet did not capture the expected cortisol awakening response. The estimated initial rate for cortisol was a 0.6% increase per minute (multiplicative factor of 1.006), but this was not statistically distinguishable from zero (p = 0.241). DHEA showed the opposite pattern, declining by 2.5% per minute (multiplicative factor of 0.975) rather than rising in the first 30 minutes. Both hormones exhibited significant declines after 30 minutes post-waking, consistent with the expected diurnal pattern: cortisol declined by 0.22% per minute (multiplicative factor of 0.998) and DHEA by 0.15% per minute (multiplicative factor of 0.9985). These point estimates are reported regardless of statistical significance to facilitate comparison with values obtained from standard hormone collection methods.

Conclusions

This study evaluated the SPIT (Saliva Procurement and Integrated Testing) booklet as a novel device for at-home collection of saliva samples to measure diurnal cortisol and DHEA patterns. Three research questions were addressed: (1) agreement between subject-recorded and electronically-recorded sampling times, (2) adherence to protocol-specified sampling times, and (3) characterization of cortisol and DHEA changes over time.

Question 1: Agreement. The Booklet and MEMs recording methods demonstrated strong agreement, with a slope near unity ($\beta = 0.995$, 95% CI: 0.981–1.009) indicating no proportional bias. A small constant bias was detected ($\beta_0 = -6.5$ minutes, $p = 0.026$), suggesting subjects recorded Booklet times slightly earlier than MEMs timestamps. This level of agreement supports the use of either recording method for research purposes, though investigators should be aware of the systematic 6-minute difference.

Question 2: Adherence. Subjects demonstrated good adherence to the +30 minute sampling window, with 87.4% of samples collected within $\pm$15 minutes of the target time. However, adherence to the +10 hour sampling window was notably lower, with only 55.0% of samples meeting the $\pm$15 minute tolerance. This discrepancy likely reflects the practical challenges of at-home collection: the +30 minute sample follows immediately after waking when subjects are focused on the protocol, whereas the +10 hour sample requires subjects to remember precise timing amid daily activities. Future studies using the SPIT booklet may benefit from reminder systems or more flexible timing windows for later-day samples.

Question 3: Diurnal Patterns. The piecewise linear mixed models revealed distinct patterns for cortisol and DHEA. At waking, mean cortisol was 7.54 nmol/L, with no significant change during the first 30 minutes (Time1 slope $p = 0.576$). Cortisol then declined significantly at a rate of 0.0092 nmol/L per minute ($p < 0.001$), equivalent to approximately 0.55 nmol/L per hour. This absence of a detectable cortisol awakening response (CAR) differs from the expected pattern and may reflect individual variability in CAR timing, the relatively sparse sampling in the first 30 minutes, or true biological variation in this healthy population. In contrast, DHEA (mean 1.91 nmol/L at waking) showed significant decline in both phases: a steeper decline from waking to 30 minutes (0.0325 nmol/L per minute, $p < 0.001$), representing approximately a 0.97 nmol/L decrease over 30 minutes, followed by a slower but continued decline (0.0009 nmol/L per minute, $p < 0.001$), equivalent to 0.05 nmol/L per hour. This continuous decline pattern for DHEA differs from the expected rise-then-fall pattern similar to cortisol, suggesting that DHEA diurnal dynamics may be distinct from cortisol in this population.

Limitations: Several limitations should be considered. First, diagnostic plots revealed some curvature in residuals for both hormone models, suggesting the piecewise linear specification may not fully capture the complexity of diurnal patterns. The Q-Q plot for DHEA showed deviation from normality in the lower tail. Additionally, the fixed knot at 30 minutes may not optimally capture individual variation in awakening response timing. The exclusion of one subject

with multiple DHEA values at the detection limit and observations with cortisol >80 nmol/L may limit generalizability.

Conclusions: The SPIT booklet device shows promise for at-home saliva collection, with strong agreement between self-reported and electronic timing methods. However, adherence to protocol timing, particularly for later samples, remains a challenge. The device successfully captured the expected diurnal decline in cortisol, though the cortisol awakening response was not statistically significant in this sample. DHEA showed a continuous decline rather than the hypothesized rise-then-fall pattern. These findings provide useful baseline data for the SPIT device and highlight considerations for future studies employing at-home salivary hormone collection.

4 Appendix

Github link:

https://github.com/tarodududu/BIOS6624/blob/main/Project%200/code.Rmd

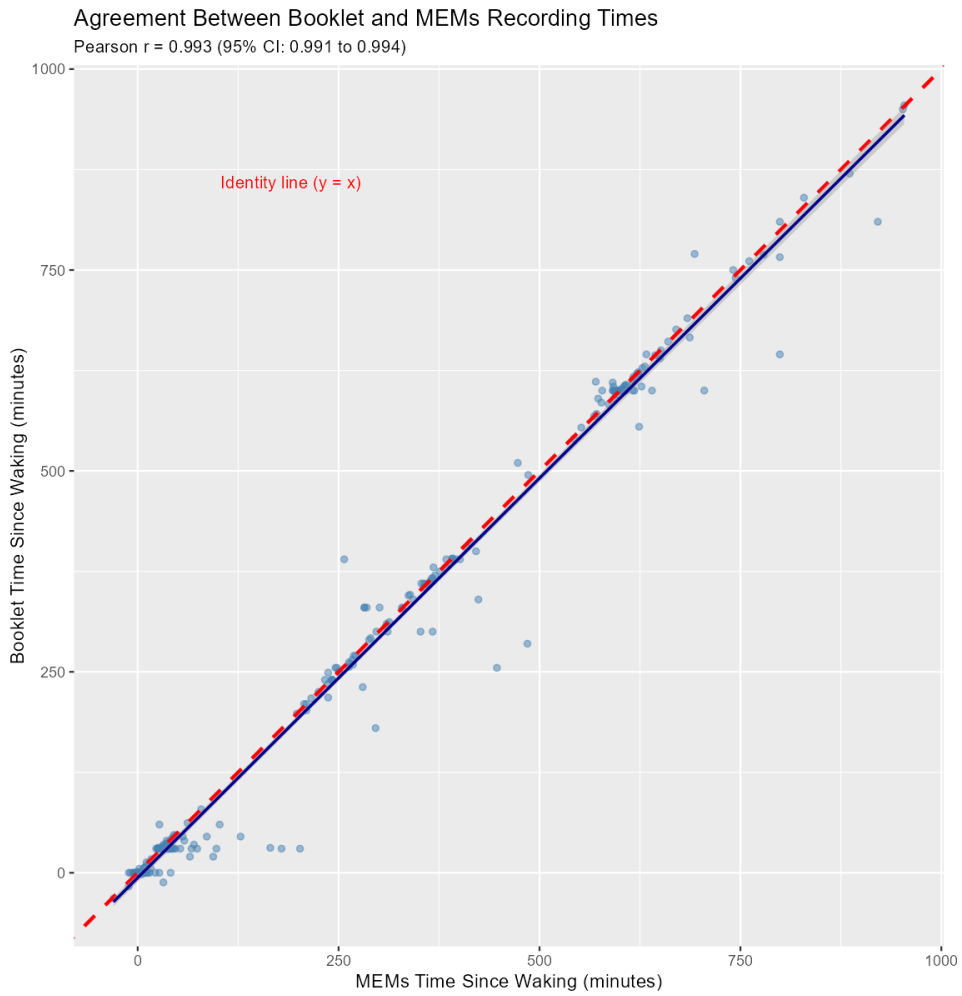

Figure 1: agreement-plot

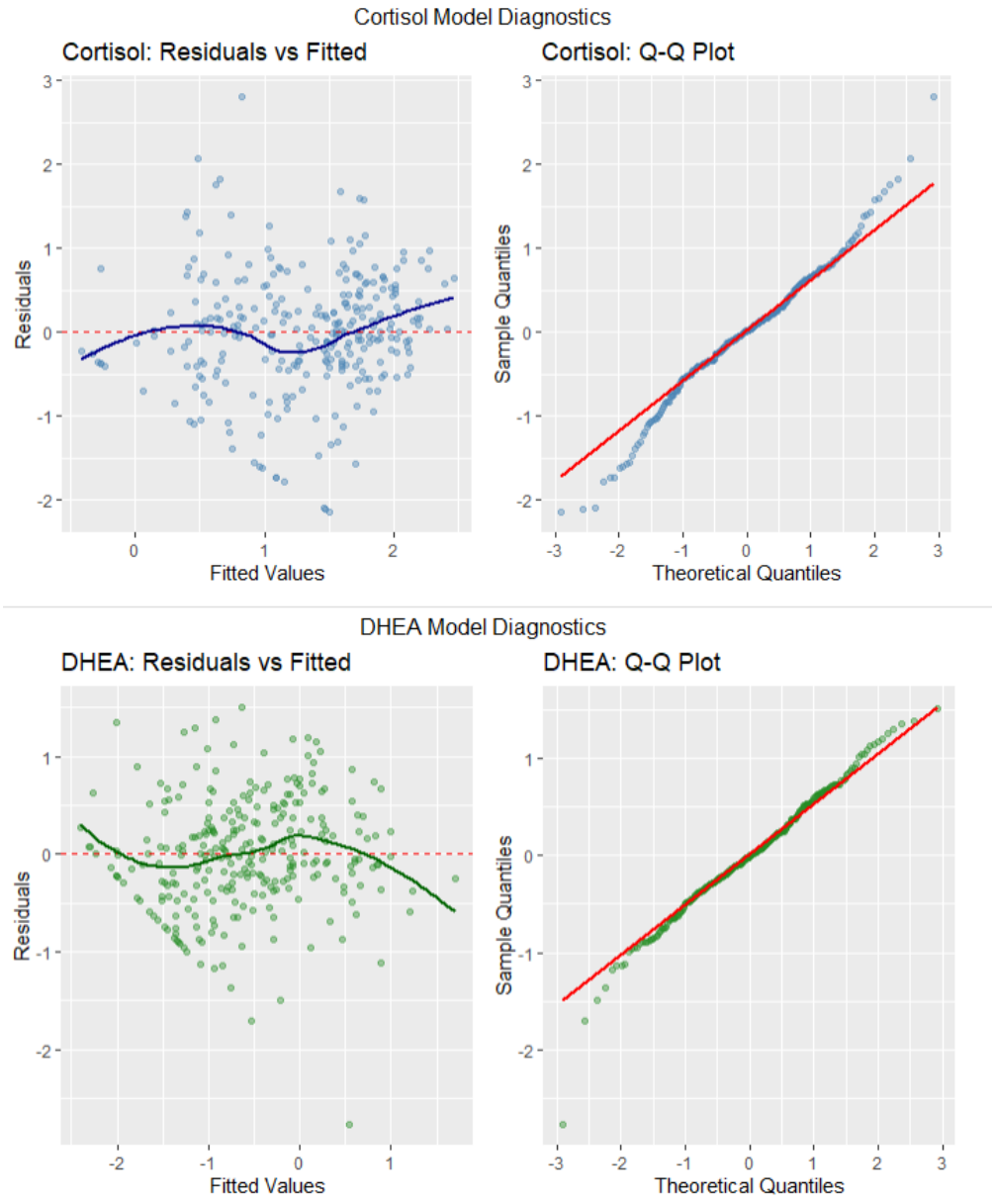

Figure 2: Cortisol and DHEA Model Diagnostics