A Comparison of Saliva & Dried Urine for Hormone Measurements

Contents
• Filter Paper Collection Device
• Creatinine Normalization
• Correlation of Dried Urine to Wet Urine
• Diurnal Rhythm of Cortisol in Urine and Saliva: Correlation?
• Relevance of Cortisol Metabolites
• Total Levels of Cortisol, Hormones and Metabolites
• Measuring Cortisol Awakening Response (CAR) with Dried Urine
• Evaluating Overnight Melatonin with a Morning Sample of Urine

Introduction
Dried urine spot testing for hormones and their metabolites is an area of recent commercial and diagnostic interest within the medical community and general public. Proponents of dried urine testing claim to provide total and diurnal rhythm type hormone measurements using a 4-sample collection regimen of urine collected then dried on filter paper. The timing of sample collection is generally set at dinnertime, bedtime, awakening and 2 hours post-awakening.

The labs who developed this testing method make claims about the clinical application of dried urine based results to free cortisol diurnal pattern, as well as aggregate data on cortisol, cortisone, sex hormones (such as estrogens) - and their metabolites - by combining samples extracted from the 4 sampling points.

On the surface, this new method appears to be an attractive manner in which to assess hormones and metabolites. However, it is an in-house lab developed test (LDT) without peer-reviewed substantiation; there is no apparent literature to support the collection methods or correlate the interpretation of these tests with specific patterns of HPA axis dysfunction or other physiological disorders.

In the interest of objectively addressing the pre-analytical methods of this testing, as well as the principle claims being made by the labs who developed the testing, the following is provided as a cursory review.

Filter Paper Collection Device
There are numerous concerns (iterated below), regarding the handling, storage, and processing of the samples in the clinical laboratory setting (1,2). This exposes the lab to a host of complications that must be eliminated in assay validation.

• There is a lack of substantial peer-reviewed studies for analysis of hormones and metabolites using filter paper. In the absence of published studies on the application of urine on filter paper
for each of the analytes being tested, doubt exists as to the validity of dried urine on filter paper for some or all of the analytes being assayed. It cannot be assumed that all analytes are equal in terms of stability; there is no known literature to support such a conclusion, while, in contrast, the stability of hormones in saliva and serum is well-established.

- The filter paper used for these labs’ dried urine hormone tests was created for blood applications, not urine. One of the lab’s test result includes the disclaimer, “The filter paper used for sample collection is designed for blood collection, so it is technically considered “research only” for urine collection. Its proper use for urine collection has been thoroughly validated.” However, no details backing this position are available in published literature. Blood has a much higher viscosity than urine (about 4X), which creates additional doubt in terms of using filter paper devices designed for blood to satisfy pre-analytical requirements.

- Soaking time is not standardized to ensure that variations in hormone and metabolite concentrations are temporal in nature and not due to differential sample uptake. Concentrations of the hormones and metabolites are likely to vary across the filter paper, with some sections of the paper having higher or lower concentrations, thereby having inconsistent storage of the analytes. Irregular measurement may result from uneven permeation of the filter paper. If the amount of urine and the size of the filter paper are not a good match, the uptake of the sample will be inconsistent in quantity across the filter paper.

- The dried urine labs instruct patients to hang their filter paper out to dry for a minimum of 24 hours, thereby leaving the samples vulnerable to the negative effects of humidity, as well as potential contamination by airborne compounds (and anything else that touches the paper, e.g. aerosol sprays, animals, smoke, or other humans).

- It is essential for any adrenal/HPA axis testing to be done on a “normal” day of stress, such as a day of work or busy errands. Since the filter paper is supposed to hang to dry without interference for 24 hours, it makes the collection at a patient’s place of work, or “on the go,” nearly impossible. The test requires that a patient be in one place, all day long, thereby negating the clinical value of the data used to gauge a patient’s ability to adapt to stress. The inability to perform this test during a typical day of stress represents a major drawback.

- Samples exposed to humidity are more at risk for fungus and other microorganisms, as well as biomolecular degradation. None of the commercially available dried urine hormone tests include desiccants to help with eliminating moisture in the sample transport container. Furthermore, the general recommendation for a location to hang the urine out to dry is in the bathroom, a humid environment, often with chemicals or other compounds suspended in the air.

There is, in general, a need for more robust, standardized protocols for sampling, storage, processing, and evaluating these collection techniques. Independent studies which support lab-developed procedures for using urine on filter paper for spot testing of hormones and their metabolites are required.

**Creatinine Normalization**

As is the case with all urinary hormone testing (most commonly associated with the 24-hour collection
method), creatinine is measured to correct for adjusting urinary hormone concentrations. While labs promoting dried urine testing make some effort to issue precautions to, and collect information from, the patient, the relatively ambiguous instructions put the accuracy of test results in doubt. Patient questionnaires lack the necessary detail established by peer-reviewed research on urine hormones (4-7).

Creatinine correction of urinary hormone values is problematic primarily because creatinine excretion is influenced by time of day, age, sex, diet, supplementation, body mass, and activity level. This variation can complicate interpretations of analytes reported in ratio to creatinine. Because creatinine is a byproduct of muscle use, its production is expected to vary with body composition and activity. As a practical example, someone who consumed alcohol and had a steak for dinner is going to have a dramatically different bedtime creatinine level than someone who had fish and water.

Additionally, although the original use of creatinine was to check the completeness of 24-hr urine specimens, many studies have shown considerable intraindividual variation in daily creatinine excretion. In one study, using 4 consecutive 24-hr urine collections, researchers observed variation ranging between 9.2% and 79.4% in the extreme values of 16 individuals. In a similar study, the scientists found a range of 63–244% in 24-hr collections from eight individuals (8,9). These large variations in recovery far exceed acceptable limits of accuracy for any testing method.

Creatinine correction is particularly questionable when applied to spot samples because creatinine excretion over short intervals also shows considerable variation. One study showed that subsequent 2-hr interval samples varied by >100% (10), and other studies have reported that spot-sample creatinine variation is several times higher than variation for 24-hr values (11,12). This inapplicability of creatinine correction to spot specimens presents a problem for hormone/metabolite testing on urine.

In addition, urine volume has been shown to be a confounder for the levels of urinary free cortisol and cortisone levels (13), where higher urine volume leads to higher levels of free cortisol in the urine. The dried urine labs do not account for total urine volume collected. Nor do they account for flow rate corrections. And there is no indication of which “catch” of urine the patient should obtain in patient instructions.

The myriad of problems associated with creatinine correction, especially on spot urine (and with no relevant research on dried urine) lend further doubt to the reliability of the dried urine testing methods.

### Correlation of Dried Urine to Wet Urine

Dried urine labs claim equivalence between dried urine and wet urine samples for the hormones tested. However, the only data provided by one of the labs is a graph (Figure 1) that plots the average of all the different hormones tested as a single data point. Since hormones can have differing stability patterns, it is inaccurate to represent that a correlation exists with a graph that shows cumulative data.
The graph is claimed to demonstrate a weighted average of four samples combined and measured for “all hormones other than cortisol and cortisone.” This averaging has globally normalized the data to minimize the variation for each hormone and appears to have limited value. Instead, showing the correlation for each hormone for fresh vs. dried urine would have been more appropriate.

There is no peer-reviewed study to date that can support the correlation between dried urine and wet urine measurement of hormones and their metabolites. The existing documentation of this correlation is authored by the labs who developed the testing. Until independent verification is performed, the idea that dried urine correlates to wet urine is unsubstantiated.

Diurnal Rhythm of Cortisol in Urine and Saliva: Correlation?

Inaccurate measurement of diurnal cortisol rhythm can lead to misdiagnosis and inappropriate treatments. The diurnal rhythm pattern of free cortisol levels is based on thousands of studies in peer-reviewed journals, and all of them use saliva as the sample source since it accurately reflects the dynamic nature of cortisol expression.

Some laboratories are attempting to expand the utility of urine samples beyond the use of 24-hour collection, including collection of urine on filter paper over 4 points over part of the day. This is supposed to allow for the analysis of a pseudo-diurnal pattern of free cortisol (shifted by 90 minutes and based on average excretion in urine over hours) and an approximation of total cortisol production calculated using cortisol metabolites. However, there are no scientific publications to support the diurnal rhythm measurements on the 4-point collection of dried urine. Unlike serum or saliva, urine cannot represent a single-point in time as excretion of up to a few hours is needed prior to collecting the next urine sample. Therefore, urine measurements are always a pooled average of several hours (or overnight) of urine excretion.

The one piece of scientific literature that was published in support of a diurnal pattern of free cortisol expression in wet urine was by Jerjes et al (2006) (14). These researchers performed urine samplings every 3 hours over the course of 24 hours. An equivalent study of free cortisol measured in urine has not been published since; it is unclear if this is due to issues with reproducibility, study funding, or concerns over the value of the data.
However, one of the proponents of dried urine hormone testing uses this study, and another study published a year prior, to show acceptability of dried urine specimens collected 4 times, from the evening before to 2 hours after awakening, combining data from the two to show a correlation between saliva and urine. They use the data from these two separate publications (15,16) and combine them (Figure 2) to show equivalence of fresh urine (collected every 3 hours over the period of a full day), to salivary cortisol collected in the same time frame. To the casual observer this is impressive, however it is fundamentally a manipulation of disparate data to arrive at a best case representation.

An additional graph was presented as a follow up (Figure 3) but no information was provided on how many patients were part of the study, nor does it show the side-by-side diurnal rhythm of the two sample types. Instead, it represents correlation data comparing three pooled saliva samples to one urine sample. They acknowledge that samples were collected at various times, and some with ACTH stimulation, which artificially increases cortisol levels. ACTH stimulation is essentially “spiking” the samples to show artificially elevated levels of cortisol whereas proper correlation data should reflect unaltered, endogenous levels of cortisol.

There are no peer-reviewed studies published to date which demonstrate a diurnal rhythm of fresh or dried urine collected at the recommended 4 time points that correlate to salivary diurnal rhythms collected at the same 4 time points.

Further research is needed to understand how to interpret such tests in the clinical setting, and what impact cortisol clearance rate (liver and kidney metabolic functions), has on these interpretations on urine. However, the acknowledgment of this opportunity is dependent upon peer-reviewed validation of the most basic aspects of the dried urine hormone testing methods and principles which have yet to be shown.

At best, the diurnal rhythm as determined with urine is a general estimate of cortisol production and timing – with numerous confounders and potential for inaccuracy and errors – while methods based on saliva and serum are real-time and correspond to the reality of a patient’s experience with situational stress and corresponding physiological functions (17).

Further, it is essential for any adrenal/HPA axis testing to be done on a “normal” day of stress, such as a day of work or busy errands. Since the filter paper device used in these tests is supposed to hang to dry without interference for 24 hours, it makes the collection at a patient’s place of work, or “on the go,” nearly impossible. The test requires that a patient be in one place, all day long, thereby negating the clinical value of the data used to gauge a patient’s ability to adapt to stress. The inability to perform this test during a typical day of stress represents a major drawback.
Relevance of Cortisol Metabolites

Proponents of dried urine hormone testing claim that having information on cortisol metabolites is valuable in assessing and treating HPA axis dysfunction. The argument is made that, without evaluating cortisol metabolites, a patient could be diagnosed as having low cortisol production, when instead the patient’s cortisol was metabolized, not making it into free circulation available for measurement. However, there is conflicting information from the few published studies looking at cortisol metabolites.

On the other hand, wet urine is a well-established sample type for determining steroidal hormone metabolite levels. But even in wet urine, the incidence of low cortisol levels corresponding to high levels of cortisol metabolites is only known to be relevant in a very small percentage of people with hypocortisolemia, due to high levels of clearance of cortisol regulated by 11 beta-HSD Type 218.

Research has uncovered some conditions that affect cortisol clearance rate or the enzymes involved in cortisol metabolism, such as extreme malnutrition and early-life stress (lower clearance), hypothyroidism (lower clearance), depression (lower clearance), and insulin resistance (higher clearance). While these data are helpful in partially explaining some anomalies between actual and expected cortisol levels, there is not enough data to understand how to evaluate the influence of cortisol clearance rate on an individual’s HPA axis function and status, and what, if any, changes to the treatment strategy should be implemented.

Also, in other literature associated with obesity, clinical depression, and chronic fatigue syndrome, free cortisol levels can vary, but metabolites of cortisol are not significantly different from healthy control populations (19-21). There is no known literature which supports a broad relevance of cortisol metabolites for the patient population (in contrast to blood and saliva hormone testing which are demonstrated to be applicable to all patients).

One dried urine lab provides a graph (Figure 4) to demonstrate agreement between free cortisol levels and metabolized cortisol levels, following this data with anecdotal information which claims that about 30% of individuals with low free cortisol have elevated levels of metabolized cortisol. However, there are several studies that show no difference in cortisol metabolites between healthy and ill patients. For example, the 2006 Jerjes paper (15) showed no difference between the levels of all cortisol and cortisone metabolites between CFS patients (who have low free urinary cortisol levels) and the healthy control population.

Additionally, the dried urine sampling regimen does not include the time point measure when there is the highest level of cortisol metabolite excretion, between noon and 3 pm, as per the peak (clock time 1200-1500) depicted in the graph labeled “Urinary cortisol metabolites” from Jerjes et al (Figure 5). This incomplete sampling reinforces the inability of clinicians to form diagnostic conclusions from dried urine results.
When examining cortisol metabolites in the context of specific disease states, evaluation of total cortisol should also be taken into account, but this is not possible using dried spot urine due to the lack of collection at the metabolic excretory peak. Physiologically speaking, between 80-90% of circulating cortisol is bound to cortisol binding globulin, and up to 10% is bound to albumin, while the rest is the biologically active (free) cortisol (23). Total (bound plus free) cortisol levels can only be measured using blood. In contrast, urinary free cortisol represents only 1% of the total cortisol secreted, while salivary testing specifically measures free, cell-signaling hormones while they are active (24).

The scientific validity of a lab test is the fundamental concern for clinicians and researchers alike. This concern is compounded when there are claims made about the value of the lab test in the context of diagnosis and treatment. Very little is known about long-term conditions of cortisol metabolism and how cortisol is cleared, rendering it inappropriate to make assumptions that low cortisol should be treated like high cortisol or vice versa, in light of metabolite levels. And in the absence of credible research or scientific literature, it should not be assumed that a patient should be treated any differently based upon the data offered by dried spot urine testing.

**Measuring Cortisol Awakening Response (CAR) with Dried Urine**

One of the most distinctive features of the 24-hour circadian rhythm of cortisol secretion is the predictable increase of cortisol that occurs in the morning, just after waking, called the cortisol awakening response (CAR). This feature is a result of two phenomena: the first is the momentum of rising cortisol levels that...
begins several hours before awakening due to normal circadian HPA axis activities (ACTH); the second, a transient (30 to 45 minute) additional increase of up to 50% in cortisol secretion due to light activation of the suprachiasmatic nucleus.

The research community has recently published guidelines required to capture the validity of the CAR measurement (Stalder et al, 2016) (25). This document was created by the leading researchers of cortisol worldwide and there is no mention of dried urine because there are no known studies supporting its validity. In these guidelines, it is imperative that the timing of the awakening and +30 min post awakening sample collections be stringently adhered to in order to make the correct clinical decisions. Taking collections more than 5 minutes later was shown to significantly alter the CAR pattern, making dried urine (which pools cortisol for 90 minutes or more) an inappropriate choice for this type of evaluation. Equally important, there is no known evidence-based research to interpret the results from urine testing's morning cortisol collections. The underlying etiology and subsequent prevalence of disease states emphasize the need to assess CAR regulation with the highest levels of precision and reproducibility, and this is only possible with saliva.

Evaluating Overnight Melatonin with a Morning Sample of Urine

Dried urine testing proponents claim that the waking level for melatonin, in the form of 6-OH melatonin sulfate, will provide useful information on how much melatonin was produced overnight. While that is uncertain due to an absence of literature on the subject, a 24-hr wet urine sample does help determine if a patient is not producing a sufficient amount of melatonin, and if they would benefit from supplementation with exogenous melatonin and/or light therapies. (26)

However, solely measuring a morning sample of melatonin may indicate low levels of melatonin and lead to the conclusion that inadequate melatonin was produced, while, in fact, that person may be phase-shifted, meaning they have a delayed dim light melatonin onset (DLMO) leading to a delayed sleep pattern. In this case, the person may make enough melatonin through the night, but if they start production later, they fall asleep later - and their morning levels of melatonin in the awakening urine sample may be low since it has not been fully metabolized. (27)

The usefulness of a waking sample is limited to individuals who are not phase-shifted and not making sufficient melatonin, but not of any use in individuals that are phase-shifted and making sufficient melatonin later in the night. These individuals may be improperly diagnosed and erroneously given higher doses of melatonin supplements which will lead to enhanced drowsiness and fatigue through the day.

In order to help overcome sleep disturbances, the most accurate and scientifically relevant approach would be a 24-hour urine collection to determine total melatonin levels produced. Correlating the 24-hour urine with salivary analysis of melatonin levels, every hour over a period of 5 to 7 hours, would more accurately establish DLMO levels of the patient. (28-30)
## Summary Comparison of Common Sample Types:

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Saliva</th>
<th>Wet Urine</th>
<th>Blood</th>
<th>Dried Urine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer reviewed studies to support testing of hormones</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Assay validation criteria met</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Maybe*</td>
</tr>
<tr>
<td>Clinical applications supported by research</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Stability of analytes demonstrated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Ability to test free, bioactive hormones</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Measures total hormone load over 24 h</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Convenient collection</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes**</td>
</tr>
<tr>
<td>Measures real-time hormone levels</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

* While the other sample types are well-established globally with common validations, dried urine’s validations remain unavailable given their LDT status.

** Urinating on paper may be convenient, however the quality of the paper and risk of contamination while the sample is drying are of concern, as explained in the text.

## Peer reviewed literature to support hormone testing in different sample types:

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Saliva</th>
<th>Wet Urine</th>
<th>Blood</th>
<th>Dried Urine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free diurnal rhythm of cortisol</td>
<td>&gt;1000 citations</td>
<td>2 citations by same group</td>
<td>not feasible</td>
<td>0</td>
</tr>
<tr>
<td>Measurement of Cortisol Awakening Response</td>
<td>&gt;450 citations</td>
<td>not feasible</td>
<td>not feasible</td>
<td>0</td>
</tr>
<tr>
<td>Melatonin for DUMO patterns</td>
<td>&gt;75 citations</td>
<td>not feasible</td>
<td>~ 39 citations</td>
<td>0</td>
</tr>
<tr>
<td>Relevant clinical applications of free cortisol</td>
<td>&gt;900 citations</td>
<td>&gt;800 citations</td>
<td>&gt; 5600 citations</td>
<td>0</td>
</tr>
<tr>
<td>Relevant clinical applications of hormone metabolites</td>
<td>not feasible</td>
<td>80 citations</td>
<td>&gt;100 citations</td>
<td>0</td>
</tr>
</tbody>
</table>
Conclusion

Given the need for reliable diagnostic tools in the clinical setting, test methods must stand up to the challenges of scrutiny. While validated methods for hormone and metabolite analysis have a well-established history in the mediums of saliva, serum, and wet urine, the dried urine testing methods lack peer-reviewed, independent substantiation and have not stood the test of time in clinical practice or research environments. Until it does, this method should be considered experimental.

References


