



METHOD ARTICLE

Validity and reliability of method used to analyse hair cortisol concentration [version 1; peer review: awaiting peer review]

Jordan Reid, Katy Parker, Lydia Clemens, Matt Bristow 

School of Psychology and Sport Science, Anglia Ruskin University, Cambridge, CB11PT, UK

V1 First published: 06 May 2021, 10:349
<https://doi.org/10.12688/f1000research.28187.1>

Latest published: 06 May 2021, 10:349
<https://doi.org/10.12688/f1000research.28187.1>

Abstract

Hair cortisol analysis is a method of analysing the stress hormone cortisol that offers great potential for helping researchers understand the long-term impact of stress and distress on the body. Hair analysis not only provides an excellent method of studying the average production of cortisol over weeks and months, but also the potential to understand cortisol levels several months before the hair was collected. Whilst research with hair samples for cortisol analysis is a fast-developing field, there has been less analysis of the methods used to analyse hair cortisol. We report two studies where the novel hair analysis method developed at the Anglia Ruskin university (ARU) Biomarker Laboratory was tested for reliability and validity. In study 1, 32 participants provided hair samples for an examination of the reproducibility of the hair cortisol analysis method. In study 2, 53 participants provided a hair sample cut from the scalp, and the methanol that the cortisol was extracted into was split between two tubes and assayed at two different laboratories with different methods (ELISA, LC-MS/MS). Overall, the results demonstrate that the methods developed to analyse hair cortisol in the ARU Biomarker Laboratory were both reliable and valid. The discussion considers further avenues for research and optimisation of the methodology.

Keywords

HCC, Hair Cortisol Concentration, Cortisol

Open Peer Review

Reviewer Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Matt Bristow (matt.bristow@aru.ac.uk)

Author roles: **Reid J:** Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Parker K:** Conceptualization, Investigation, Methodology, Project Administration, Resources, Writing – Review & Editing; **Clemens L:** Investigation, Project Administration, Writing – Review & Editing; **Bristow M:** Conceptualization, Formal Analysis, Methodology, Project Administration, Supervision, Validation, Writing – Original Draft Preparation

Competing interests: One of the authors (MB) is the Director of the ARU Biomarker Laboratory and one (KP) is the ARU Biomarker Laboratory manager. The laboratory is a research and commercial laboratory within Anglia Ruskin University. The laboratory undertakes research with biomarkers within Anglia Ruskin University and provides commercial hair and saliva testing services.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Reid J *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Reid J, Parker K, Clemens L and Bristow M. **Validity and reliability of method used to analyse hair cortisol concentration [version 1; peer review: awaiting peer review]** F1000Research 2021, 10:349 <https://doi.org/10.12688/f1000research.28187.1>

First published: 06 May 2021, 10:349 <https://doi.org/10.12688/f1000research.28187.1>

Introduction

Cortisol is a hormone secreted by the zona fasciculata of the adrenal cortex which binds to the glucocorticoid receptor (GR) found in almost every vertebrate cell. The ubiquity of the GR means that cortisol's effects across the body are diverse and include metabolic (increased blood sugar production), immunological (such as reduced resistance to certain infections), circulatory (acute elevation of blood pressure), renal and central nervous system (such as altered mood and reduced libido) effects (Becker, 2001). Cortisol secretion follows a now well-defined circadian rhythm (Clow *et al.*, 2004; Hucklebridge *et al.*, 2005) and cortisol secretion can also be substantially increased in response to psychological or physiological stress (Dickerson & Kemeny, 2004). Cortisol's wide influence over cellular function means that stress induced changes in cortisol have the potential for substantial changes in physiology, behaviour and health.

Research exploring psychosocial variables in relation to cortisol secretion have predominantly measured cortisol in saliva, blood or urine samples, which allow for the measurement of near-current levels of cortisol secretion (saliva and blood) and recent levels (urine). Whilst such research has proved hugely useful, not all research questions seek to explore near-current or recent cortisol output. Researchers who want to explore how personality, traumatic events, or mental health conditions, for example, impact on cortisol secretion and subsequent health may well find a measure of longer-term cortisol secretion to be very useful. The assessment of average cortisol activity over weeks and months is challenging, with the traditional measures from serum, saliva or urine as it requires sampling over many days with multiple samples a day to capture the full complexity of the circadian variation. In addition, the traditional measures of cortisol can only be used prospectively, whilst analysis of hair for cortisol offers the tantalising possibility of retrospective cortisol measurement. For researchers wanting to understand cortisol secretion over longer time periods, with the potential to measure cortisol secretion that occurred before the research had even started, the analysis of hair samples may well provide the solution.

Analysis of hair samples for measuring cortisol levels was first reported in the behavioural literature in 2004 (Raul *et al.*, 2004) and since then there has been a rapid increase of research using measures of hair cortisol to explore stress and or the reaction of the body to challenge in some form. A recent meta-analysis (Stalder *et al.*, 2017) identified 72 papers exploring some aspect of chronic stress with hair cortisol. There is a wide variety of research being undertaken. For example, higher hair cortisol in comparison to control groups has been found in unemployed individuals (Dettenborn *et al.*, 2010) shift workers (Manenschijs *et al.*, 2011), endurance athletes (Skoluda *et al.*, 2012), chronic pain patients (Van Uum *et al.*, 2008), alcoholics in withdrawal (Stalder *et al.*, 2010), undergraduates students during term time (Stetler & Guinn, 2020), dementia caregivers (Stalder *et al.*, 2014), people with post-traumatic stress disorder (Steudte *et al.*, 2011) and in children with early life adversity (Aas *et al.*, 2019; Bhopal *et al.*, 2019).

Underpinning the burgeoning use of hair cortisol in psychological research is the methodology used to determine the concentration of hair cortisol found in the hair sample. There is no single standardised approach at present for the measurement of hair cortisol but laboratories largely use a similar sequence of steps: the hair is washed to remove contaminants from the surface of the hair, the washed hair is dried, the sample is then cut or ground to increase the surface area of the hair available to extract cortisol from, the cortisol is extracted into a solvent (usually methanol over a period of time), the methanol and hair are separated, and the methanol is evaporated to dryness and analysed for cortisol concentration. The choice of assay method varies considerably but the commonest methods used are liquid chromatography with tandem mass spectrometry (LC-MS/MS) or an antibody-based assay, commonly enzyme-linked immunosorbent assay (ELISA).

The methodology developed by our laboratory was intended to be a scalable method of hair analysis that could be undertaken for large numbers of hair samples and had the minimum number of steps to reduce the chance of error and reduce costs. Our method reduces the number of steps by not transferring the sample to another container; it is kept in the same tube from initial weighing through to the end of the methanol incubation. The procedure is also very scalable. The 'rate-limiting' step for our hair analysis procedure is the grinding of the hair and this has been partially automated and is highly scalable. Final quantification of cortisol is performed via an ELISA assay, an assay which is rapid and cost effective.

In this paper we describe the methodology developed by the Anglia Ruskin university (ARU) Biomarker Laboratory and two studies that evaluate how reproducible our method is and the extent to which our assay results relate to LC-MS/MS analysis performed at another laboratory.

Aims

Study 1: Assay reproducibility

This study aimed to test the reproducibility of the hair cortisol assay and to develop a protocol that allowed the laboratory team to generate a number of aliquots of hair that had the same or very similar levels of hair cortisol.

Study 2: Assay validity

Study 1 demonstrated that the hair assay procedure developed at the ARU Biomarker laboratory was highly reliable, but it did not demonstrate that the method creates valid results. To test this key hypothesis, we split the methanol, into which the hair cortisol extracts, into two tubes rather than the standard procedure of using one tube. One of the pair of tubes was analysed for cortisol via LC-MS/MS at Dresden Labservice GmbH, a world leader in hair cortisol testing, and one via the ELISA in the ARU Biomarker Laboratory, as per study 1.

Methods

Study 1: Assay reproducibility

Participants

A total of 30 male and two female participants aged between 18 and 55 years old were recruited to participate in this study using an opportunity sample of people having their hair cut in a specific barbers in Cambridge, UK. Hair samples were collected over multiple visits to barbers over eight days. The Participants were included if their hair was 3 cm or less in length and they were Caucasian. There were no other inclusion or exclusion criteria. Recruitment was stopped once 32 samples had been recruited as this number would enable us to fit all of the replicate A and B samples on the same ELISA plate, thus removing any inter assay variability from the analysis. In addition, 32 samples would give us substantial power (.99) to detect a correlation greater than $r = .65$, should such a relationship exist between the two variables. Given that replicates A and B should correlate very highly this was considered to be an adequate sample size.

Design

There were three linked experiments in study 1.

- (1) Reproducibility: From each participant, two aliquots were analysed for hair cortisol with the aim of testing the reproducibility of the assay.
- (2) Incubation times: Hair samples from each participant were tested using the hair cortisol method after 3, 21, 24, 28 and 48 hours of methanol incubation. This was intended as an initial test to identify where the hair cortisol concentration (HCC) plateaued during the methanol incubation.
- (3) Grinding media: Hair samples from each participant were analysed using the standard hair cortisol methodology using either three or five ceramic grinding balls.

Hair retrieval protocol

Hair was collected from volunteers from people who were having their hair cut at local barbers. Participants were given information about the research after arrival at the barber shop and those who agreed to participate were asked to complete a written informed consent form and a brief hair questionnaire, which asked for specific detail of their hair treatment, for example, what styling products were used (this data is not reported further).

Once the barber was in place and ready to begin shaving, one of the authors (JR) held a small, rectangular plastic container approximately 15 centimetres below the barber's clippers, to collect the shaved hair from temporal and occipital regions of the participant's head. The collected hair was stored in a tin foil 'envelope', so the hair was stored securely but not hermetically sealed, and the samples were returned to the ARU Biomarker laboratory for storage at room temperature.

To create aliquots of the hair samples for analysis the participants' samples were mixed using two clean tweezers for five minutes to ensure a thorough mixing of the hair. The aliquots were weighed using a precision balance. The hair from each participant was aliquoted in 25 mg amounts into 4.5 ml PPCO vials that are used from the weighing stage through to the methanol incubation.

Hair analysis protocol

Hair was analysed using the method reported by [Parker & Bristow \(2020\)](#). 25mg samples of hair in the 4.5-ml PPCO vials were each washed twice in 2 ml of isopropanol to remove external contaminants. The isopropanol was then removed from the vial and the hair allowed to dry in a microbiological safety cabinet for a minimum of 48 hours until fully dry. Once fully dry, five ceramic balls (Lysing Matrix M, MP Biomedicals, LLC) were added to each tube and the hair samples

ground to a powder using Fast Prep-24¹ (MP Biomedicals, LLC). To extract cortisol, we added 2 ml of methanol to each sample and incubated the samples for 24 hours at room temperature rotating the samples constantly.

The hair, methanol and ceramic balls were decanted into a polypropylene tube (product code: 51.1534, Sarstedt AG & Co, Germany) and centrifuged at 1500 RCF to separate the ceramic balls from the rest of the mixture. The tube was then centrifuged at 3000 RCF to separate the ground hair and methanol, and 1.4 ml of the clear methanol supernatant was decanted into a 2-ml polypropylene microtube (product code: 525-0539, VWR). The supernatant microtubes were then centrifuged to dryness in vacuum centrifuge (Scan Speed 40, Labgene) at 37°C and 1700 RPM for three hours, and then frozen at -80°C until required for the cortisol ELISA. Cortisol levels were determined using a commercially available competitive ELISA (product code: #1-3002; Salimetrics, US).

Samples were thawed and reconstituted with Salimetrics' cortisol assay diluent and the samples were then assayed in accordance with the manufacturer's protocol. A Salimetrics high and low control sample run in each ELISA plate run. The 32 samples for HCC replicates A & B were run in singlet to ensure that they were all run on the same assay plate initially. All other samples from the study were run in duplicate. Samples were re-run in duplicate at a higher dilution if the ELISA cortisol concentration was higher than 3 µg.dl⁻¹. Replicate repeats were accepted as cortisol concentration coefficient of variation for the duplicates was less than 15% (unless the absolute difference between repeats was within 0.03 µg.dl⁻¹). The samples were run over 9 plates with an inter assay coefficient of variation for the high control of 3.62% and 3.08% for the low control. The intra assay coefficient of variance was 3.77%. The results were expressed as the picograms of cortisol per milligram of hair.

Study 2: Assay validity

Participants

The participants in study two were separate from those in study one. Staff and students in the ARU Department of Psychology were invited to participate in this study by email and advertisements in lectures and were offered £2.50 or 0.5 research credit for participation. We expected to find a strong correlation between ELISA and LC-MS/MS and the power calculation study 1 remained the same that we wanted a minimum of 32 participants to test the correlation between ELISA and LC-MS/MS and that this would give us substantial power to detect a correlation of greater than 0.65, should that exist. Fifty-three people (eight males, 45 females and one person who declined to answer) with a mean age of 20.98 years (SD: 4.15) provided a hair sample for analysis. The only inclusion criteria was having hair at least 1cm long. One person accidentally provided two separate hair samples one week apart, so the first hair sample provided was used in the analysis and the other excluded from analysis. All participants gave written informed consent to participate.

Design and analysis

A hair sample (see hair collection, below) was taken from each participant along with a brief demographic questionnaire that asked questions about age, gender and a range of questions about the participants height, weight and hair washing methods that are not reported further. The ARU Biomarker Laboratory extracted the cortisol from the first 1 cm of hair from the scalp into methanol using the methods detailed in study 1. The methanol from each sample was then split equally into two polypropylene tubes (0.7 ml per tube) and dried down using the standard protocol. The tubes were then stored at -80°C. One set of tubes was then sent to Dresden Labservice GmbH for analysis via LC-MS/MS (see [Gao et al., 2013](#)) and the other set was analysed via ELISA at the ARU Biomarker Laboratory using the same methods detailed in study 1. All samples were analysed in duplicate. The samples were run over 2 plates with an inter assay coefficient of variation for the high control of 9.75% and 10.16% for the low control. The intra assay coefficient of variance was 9.64%. The results were expressed as the picograms of cortisol per milligram of hair.

Hair collection

Hair samples were collected from each participant from the posterior vertex of the scalp as this area has the lowest coefficient of variation compared to hair sampled from other areas of the scalp ([Sauvé et al., 2007](#)). The hair sample size was approximately a five mm diameter of hair in total and the samples were taken from two different locations of the posterior vertex to reduce the visible signs that hair had been cut. The hair was tied with linen thread as close to the scalp as possible and the hair was then cut as close to the scalp as possible using clean scissors and stored in a tin foil envelope at room temperature.

¹The grinding settings are: 6 m/s for 30 seconds × 6 with 5 min cool down between grinds.

Ethics

The research was approved by the Faculty of Science & Engineering Faculty Research Ethics Panel at Anglia Ruskin University (Study 1 ethics number FST/FREP/18/805; Study 2 ethics identifier: FST/FREP/13/378). All participants provided signed informed consent before participation.

Statistical analysis

Analysis was conducted using **Jamovi** 1.63 (The Jamovi Project, 2020) and IBM **SPSS** version 26 (IBM SPSS, 2019) with $p < .05$ considered as a significant difference. Power Analysis used **G*Power** 3 (Faul *et al.*, 2007). For all HCC variables in study 1 there were three extreme outliers (see **Figure 1A**) that were more than three times higher than the interquartile range from Tukey’s hinge at the 75th centile, and a further two points that were outliers for most variables. (**Figure 1B** shows the relationship between replicates with the five outliers removed.) In addition, the hair cortisol variables show a significant positive skew and kurtosis. The HCC data in Study 2 also contained a number of outliers with a positive skew and significant deviation from normal distribution. **Table 1** shows descriptive statistics for the full datasets of both study one and two.

Many researchers working with hair cortisol use log transformations to reduce the skew seen with hair cortisol and to reduce the impact of high outlying values that are often found. We have not used log transformations in our analysis, partly because of theoretical objections that researchers have raised to log transformations (Feng *et al.*, 2014) and on a practical level because log transformations did not remove the distribution problems observed in our data. We opted instead to use non-parametric analysis and to provide supplemental parametric analysis where it may be helpful to compare with previous research. Where parametric analysis is used it is noted along with the outliers removed.

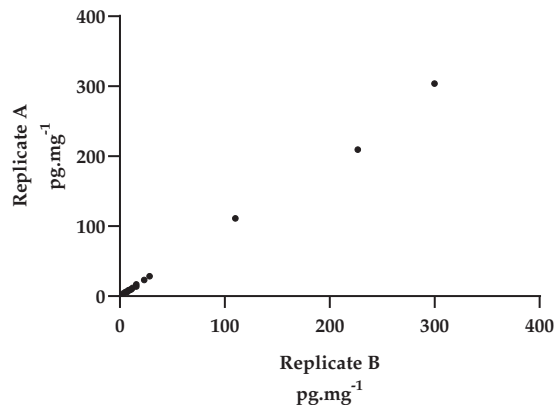


Figure 1A. Hair cortisol concentration concentrations from two samples of hair from the same individuals (all participants included).

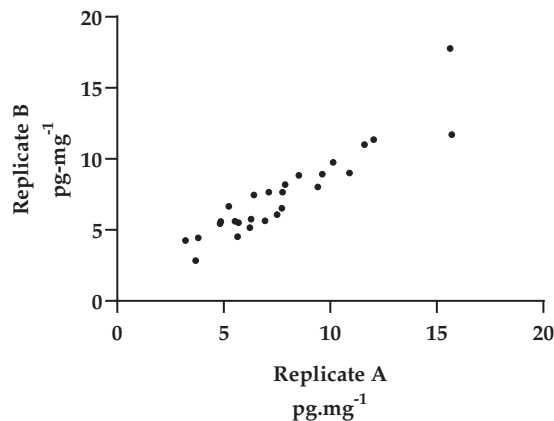


Figure 1B. Hair cortisol concentration concentrations from two samples of hair from the same individuals (Five outlying samples removed).

Table 1. Descriptive statistics for the hair cortisol concentrations for the complete data sets in study one and study two.

	N	Mean pg.mg ⁻¹	Median pg.mg ⁻¹	SD pg.mg ⁻¹	Minimum pg.mg ⁻¹	Maximum pg.mg ⁻¹	Skewness	Kurtosis	Shapiro-Wilk p
<i>Study One</i>									
Replicates									
Replicate A	32	28.07	7.75	65.14	3.21	299.97	3.52	12.07	< .001
Replicate B	32	27.04	7.66	63.24	2.84	307.85	3.68	13.86	< .001
Incubation test									
3 hrs	25	29.38	7.84	57.82	4.69	245.62	2.97	8.61	< .001
21 hrs	25	29.08	8.56	56.03	4.72	207.42	2.73	6.40	< .001
24 hrs	25	28.78	7.56	57.88	4.85	225.28	2.89	7.49	< .001
28 hrs	25	30.47	8.36	65.28	3.80	299.73	3.48	12.77	< .001
48 hrs	25	31.86	8.38	65.55	5.26	283.23	3.13	9.73	< .001
Ceramic ball test									
3 balls (24 hours)	26	29.42	8.06	60.5	4.03	225.89	2.85	7.15	< .001
<i>Study Two</i>									
ELISA	53	7.73	6.41	6.28	0.98	41.11	3.27	14.84	< .001
LC-MS/MS	53	5.96	3.84	5.7	1.30	36.70	3.47	16.07	< .001

SD = standard deviation; ELISA = enzyme-linked immunosorbent assay; LC-MS/MS = liquid chromatography with tandem mass spectrometry.

In both studies Bland-Altman plots (Bland & Altman, 1986) were used to analyse the *agreement* between the methods rather than the *association* (assessed via correlation). When Miller, Plessow, Rauh, Gröschl, & Kirschbaum (2013) compared the ELISA and LC-MS/MS results with salivary analysis, they reported high correlations between the two methods but that all ELISA results were higher than the LC-MS/MS method (ranging between 27.9% and 259.3% higher depending on the ELISA assay used). It is therefore not our expectation that the LC-MS/MS and ELISA will agree in absolute terms but rather than the difference between the two methods will be consistent over the range measured.

Spearman's Rho non-parametric correlation was used to measure the association between the HCC replicates A and B (Study 1) and ELISA and LC-MS/MS HCC from the same hair cortisol extract (Study 2) and Friedman's non-parametric test for repeated measures used to test for a difference between the hair cortisol replicates A & B in study 1 and between ELISA and LC-MS/MS analysis in study 2. Pearson's Product Moment correlation and Student's t-test have been used to provide a parametric version of the above analysis to allow comparison with other research. Friedman's repeated measures non-parametric test was used to test whether the HCC different across the five different incubation times for hair and between samples ground with 3 ceramic balls or 5 ceramic balls. Where there were more than two levels Durbin-Conover pairwise tests used to identify the location of any differences.

Results

Study 1: Assay reproducibility

There was a very significant correlation between the two replicates of hair cortisol using the full dataset ($r_s = .967$, $N = 32$, $p < .001$; $r = .995$, $N = 32$, $p < .001$, see Figure 1A) (Reid, Parker & Bristow, 2020). Friedman's test revealed no significant difference between the two replicates ($\chi^2 = .50$, $df = 1$, $p = ns$). As noted above, parametric analysis on this dataset is unreliable given the distribution of the data but to allow a better comparison with other published work we have included some parametric analysis. With the five outliers the Shapiro-Wilk test for normality of data remains significant indicating that the data is not normally distributed. The Pearson's product moment correlation between the replicates remains highly significant ($r = .926$, $N = 27$, $p < .001$; $r_s = .945$, $N = 27$, $p < .001$) and there was no significant difference in the means for replicate A (7.78 mg.pg^{-1} , $SD = 3.27$) and replicate B (7.46 pg.mg^{-1} , $SD = 3.06$) ($t = 1.35$, $df = 26$, $p < .190$). The Bland-Altman plot (see Figure 2 below) suggests a good agreement across the range of HCC means though two points fall on or outside the ± 1.96 lines of agreement.

The Friedman non-parametric repeated measures test indicated a significant difference between five incubation times (3, 21, 24 and 48 hours; $\chi^2 = 10.9$, $df = 4$, $p = .027$). Durbin-Conover pairwise comparisons showed that the 24 h HCC were significantly lower than the other levels but there were no other significant differences. The Friedman n test revealed no significant difference between using three or five ceramic balls ($\chi^2 = 1.00$, $df = 1$, $p = ns$) to grind the hair samples.

Study 2: Assay validity

The LC-MS/MS hair cortisol concentrations were significantly lower than the ELISA concentration ($\chi^2 = 32.7$, $df = 1$, $p < .001$; see Table 1) but there was a very strong correlation between the LC-MS/MS and ELISA hair concentrations

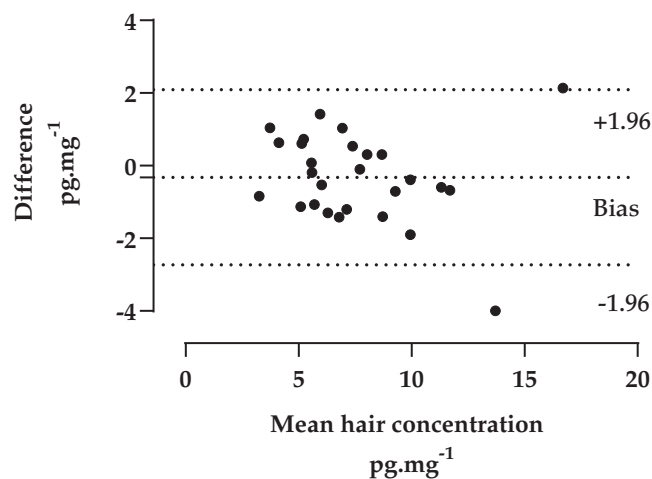


Figure 2. Bland-Altman plot reporting the differences between hair cortisol replicates (Replicate B – Replicate A) plotted against the mean replicate concentration. Five outliers have been removed from the data. The top and bottom lines indicate + or – 1.96 standard deviations from the replicate differences.

($r_s = .905$, $n = 53$, $p < .001$; $r = .911$, $n = 53$, $p < .001$) (Bristow, Clemens & Parker, 2020). The scatterplot showed a strong linear relationship with one discrepant value (see Figure 3). The Bland-Altman plot (see Figure 4) also clearly indicated one value with a highly anomalous result but otherwise the values fell within the boundaries.

Discussion

Our initial experiment in Study 1 examined how well two aliquots of hair from the same person yielded the same hair cortisol concentration. This experiment tested two interlinked research questions: 1) were the aliquots of hair that had been created very similar for hair cortisol levels? And 2) were our assay results reproducible? We found a very high correlation between the two aliquots of hair and this supports both questions with no evidence of difference between hair cortisol levels found in the two aliquots. In addition, the average coefficient of variance for the HCC of the two replicates was 8.51% (SD: 6.2%), which again shows a close agreement between the replicates. Correlations are not the ideal way to measure agreement between two laboratory measures, and the Bland-Altman plot can be of more use (Grouven *et al.*, 2007). The Bland-Altman plot (see Figure 2) shows that all but two of the datapoints were within the ± 1.96 limits of agreement and there is no pattern in the plotted data that would suggest any systematic bias. There are two points that lie outside the limits of agreement, one high and one low, and these both came from samples towards the top of the range of concentrations reported.

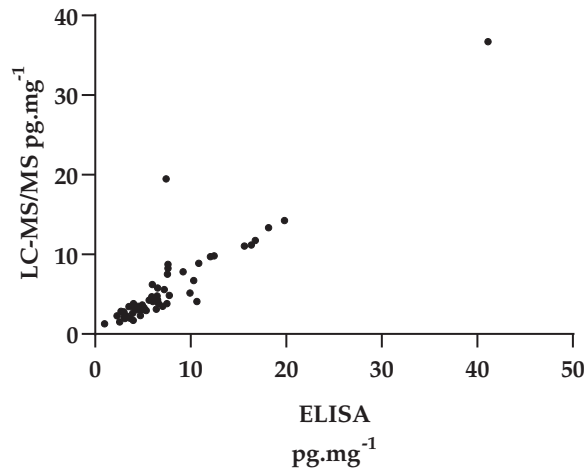


Figure 3. Scatterplot of the analysis of hair cortisol extract analyzed by ELISA and LC-MS/MS.

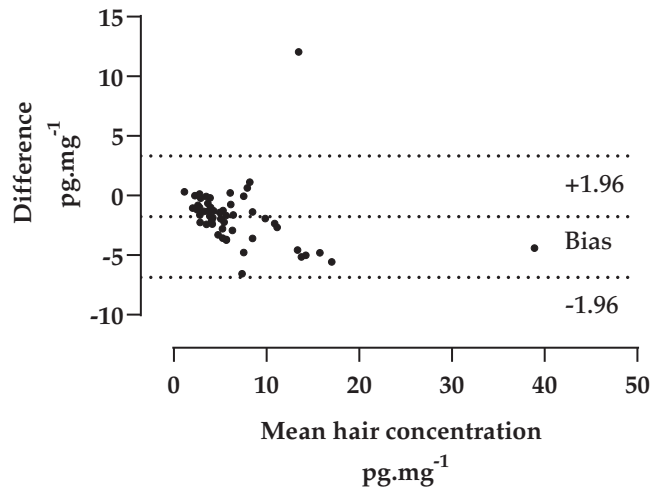


Figure 4. Bland-Altman plot reporting the differences between hair cortisol replicates in study 2 (LC-MS/MS - ELISA) plotted against the mean replicate concentration. The top and bottom lines indicate + or - 1.96 standard deviations from the replicate differences.

Our results showed that the error was being kept to a low level and the results were reproducible. This supports both our research questions. We can be confident that the aliquots of hair created have similar cortisol levels in each aliquot and that our assay is providing a reproducible result with the current protocol.

Having established that the aliquots were very similar in hair cortisol levels and that our assay showed high reproducibility, we wanted to explore whether the incubation times we are currently using are optimal and whether the grind quality could be reduced without impacting performance.

The incubation test was intended to be the first of a series of tests to understand the time of peak extraction of cortisol and understand the variability of extraction around the peak extraction point. We might have expected that incubating for three hours would have extracted less cortisol because of the relatively short extraction time and the timings we chose for this first test were aimed at giving us more information around the 24-hour time, as this is the time that we currently use. It is of interest for us at what point the cortisol extraction into methanol peaks and plateaus. Though 24 hours provides a highly reproducible result, it may be that we can optimise the assay further with a longer incubation or that we can reduce the incubation times, which may allow for faster analysis of samples. Unexpectedly, we found that hair cortisol levels were stable over the incubation times except that the 24-hour incubation was slightly, but significantly, lower than the other incubation times. It is of interest that HCC shows no sign of increasing with incubation and then plateauing. Instead we have – save for a very small dip at 24 hours – the same level of extraction over the time. This means that the 24-hour incubation period is not necessary for maximal extraction and we could likely achieve the same results in three hours, or even shorter. It also likely means that the hair grind level could be reduced. We currently grind to a very fine grind which exposes a substantial surface area for methanol to extract cortisol from and appears to allow for a fast extraction time. Reducing the number of ceramic balls from five to three does visually reduce the fineness of the grind (unpublished observations from our laboratory) but we see no impact on the HCC extracted at 24 hours. Aside from reducing laboratory testing costs this suggests that research should look at reducing the degree of grinding and how this impacts on extraction of HCC. Grinding clearly damages the hair and it may be that we can achieve the same or better extraction of hair cortisol with less grinding.

We can find no reason for this slight fall in HCC at 24 hours. The samples were all assayed at the same period and the samples were randomly distributed over the ELISA plates used. There is always a minor variation between ELISA plates: had we had allocated all of the 24-hour incubation samples to one plate alone and the other conditions to different ELISA plates, this might have accounted for this difference, however this did not happen and the 24-hour samples were randomly allocated over plates. There was no difference in the hair weights for the various levels of incubation ($F_{4,96} = 2.30$, $p = ns$). Indeed, the hair weights were consistent with a range of 24.5 mg to 25 mg across all the conditions. There is no reason the samples would extract slightly less at 24 hours than they would at 21 hours or 28 hours (or indeed three hours or 48 hours) other than random chance or a small systematic error in the laboratory work on the 24-hour samples which may have been processed separately to the other incubation times. For example, the 24-hour aliquots of hair may be the ones that had, by chance, the slightly lower levels of hair cortisol in for that participant or that the researcher chose a different pipette to aspirate off the supernatant than selected for the other conditions, and whilst both pipettes are calibrated one may pipette slightly less than the other. In normal laboratory work, the small differences would not be observed, but given the nature of this experiment and the high reproducibility of the assay, an effect may be observed. The incubation test needs to be repeated in conjunction with varying grinds of hair to further optimise the assay but the slight drop at 24-hours does not create any issue for our standard assay protocol as all of our analysis is currently carried out with a 24-hour methanol incubation.

Study 1 supports the view that the hair cortisol assay provides highly reproducible results, something that is essential for any well-functioning assay. The other key aspect of any assay is whether the results obtained measure what they claim to measure.

The results from study 2 showed that ELISA method significantly over reported HCC with the median average for ELISA $2.57 \text{ pg} \cdot \text{mg}^{-1}$ (66.9%) higher than the median for LC-MS/MS. This inflation of the cortisol results is not trivial but is also not unexpected based on the previously mentioned finding from [Miller *et al.* \(2013\)](#) that LC-MS/MS measures of salivary cortisol were significantly lower than those found by ELISA. There was very good correlation between the LC-MS/MS and ELISA measurement of cortisol extracted from the hair samples in study 2 and is comparable to correlations reported between saliva and LC-MS/MS by [Miller *et al.*](#)

The Bland-Altman plot (see [Figure 3](#)) showed no clear systematic pattern in the differences with the caveat that as in study 1 the distribution was substantially positively skewed with most values falling at the lower end but with a tail of large values.

There is one very discrepant value in the results where LC-MS/MS recorded 19.48 pg.mg^{-1} compared with 7.44 pg.mg^{-1} from the ELISA method, odd since ELISA tended to over report hair cortisol in this study. Analysis via LC-MS/MS or ELISA is not perfect and there is always a small chance of an undetected technical issue, operator error or accidental contamination. This discrepant value may also, of course, reveal a potential confound which is causing one of the two assays types to give a false reading. We recorded a wide range of information about how people treated their hair, and it is perhaps noteworthy that the participant whose hair showed the discrepant values had dyed their hair the night before it was cut. Only one other participant in our study had dyed their hair so close to having their hair sample taken. This participant had accidentally been allowed to enrol in the study twice and had given a hair sample a week before they dyed their hair and then a second time on the day their hair was dyed. We had excluded this second sample from analysis as we could not have two samples from the same person. However, there is no indication from this excluded sample of any additional discrepancy between the ELISA measure (7.73 pg.mg^{-1}) or the LC-MS/MS measure (4.39 pg.mg^{-1}). This is a 75.9% inflation by the ELISA, which is not far from the average median inflation found in by the ELISA method in the study of 66.9%.

By splitting the extracted methanol between ELISA and LC-MS/MS we are explicitly testing the measurement of HCC in the extract rather than testing whether given samples of the same hair, both laboratories would arrive at the same result. Our results do show that with ELISA analysis of the extract our methodology produces a good correlation with LC-MS/MS analysis of the same extract. Testing the relationship between two different hair analysis methods from grinding through to analysis by ELISA or LC-MS/MS would be useful but any differences in the results could be due to the hair extraction methodology used in the other laboratory rather than the end assay used (ELISA vs LC-MS/MS).

In this article we report the methodology used by the ARU Biomarker Laboratory to analyse hair cortisol levels via ELISA and two studies used to test the reproducibility and validity of the assay. The assays show high reproducibility and high validity, and the data suggest the methodology is fit for purpose. The methods used in study 1 to create a series of aliquots of the same participant's hair was successful in allowing us to have access to multiple samples of each person's hair. This allows for a range of validation testing to be carried out on the same hair samples.

There is considerable scope for further optimisation work for the assay. The incubation times can be usefully analysed along with different degrees of grind from uncut hair through rough cut to fine powder to explore the lowest degree of grind that can produce reproducible results. Our initial development work (unpublished observations) suggested that 25 mg of hair provided reproducible results but that a range of around 10 mg to 100 mg of hair could be used to produce reproducible results. Outside these values there were concerns about the reproducibility. A more systematic evaluation of this could help researchers better understand the amount of hair required for reliable testing and the reliability of analyses conducted on samples outside of this range.

The research presented here confirms the hair analysis methodology developed at ARU Biomarker laboratory is both a reproducible and valid measure of hair cortisol. This allows for a scalable method for analysing cortisol levels from hair that can be undertaken for large numbers of hair samples while minimising the number of steps to reduce chance of error and costs. The findings evidence the reproducibility and validity of hair cortisol analysis as a measure of psychological and physiological stress. This type of work is essential in underpinning the application of hair cortisol measurement to increase our understanding of longer-term impacts of psychological stress on the human body.

Data availability

Underlying data

Figshare: Hair cortisol concentration method validation study 1. <https://doi.org/10.6084/m9.figshare.13352822.v3> (Reid, Parker & Bristow, 2020)

This project contains the following underlying data:

- Study 1 hair validation Reid repository in.sav and.csv formats (The files contain the hair cortisol concentrations (HCC), the underlying ELISA measurement of cortisol (not corrected for hair weight) and hair weights for each condition in study 1 (replicate A & B, incubation over 3, 21, 24, 28 & 48 hours and a 3 ceramic ball grind condition))
- Study 1 Reid detailed ELISA data.csv (This file contains the detailed ELISA (Enzyme Linked Immunosorbent Assay) records for all of the hair cortisol data. This includes replicate data for cortisol concentration, optical density (OD; 450nm) and difference optical density (620nm), the coefficient of variation for cortisol and difference optical density replicates and the sample dilution.)

Figshare: Hair cortisol concentration method validation study 2. <https://doi.org/10.6084/m9.figshare.13359695.v3> (Bristow, Clemens & Parker, 2020).

This project contains the following underlying data:

- Study 2 ELISA LCMSMS Reid repositoryv2.sav
- Study 2 LCMSMS Reid.csv (The .sav and .csv files contain the hair cortisol concentrations (HCC), the underlying ELISA measurement of cortisol (not corrected for hair weight) and hair weights for the samples analysed by both LC-MS/MS and ELISA. In addition, the files details which ELISA plate the samples were processed on and the order in which the aliquots of methanol went into the tubes (1 = ELISA tube first and 2 = LCMS/MS tube first). Age, BMI, height, weight, medication, hair colour and hair washing variables have been removed as they may allow identification of participants.)
- Study 2 Reid ELISA data.csv (This file contains the detailed ELISA (Enzyme Linked Immunosorbent Assay) records for all of the hair cortisol data. This includes replicate data for cortisol concentration, optical density (OD; 450 nm) and difference optical density (620 nm), the coefficient of variation for cortisol and difference optical density replicates and the sample dilution. There is a sample number to identify which samples came from the same participant but this a randomised code and does not relate to the original participant numbers.

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) (CC-BY 4.0).

References

- Aas M, Pizzagalli DA, Laskemoen JF, *et al.*: **Elevated hair cortisol is associated with childhood maltreatment and cognitive impairment in schizophrenia and in bipolar disorders.** *Schizophr Res.* 2019; **213**: 65–71.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Becker KL: *Principles and Practice of Endocrinology and Metabolism*. 3rd ed. Lippincott: Williams & Wilkins; 2001.
- Bhopal S, Verma D, Roy R, *et al.*: **The contribution of childhood adversity to cortisol measures of early life stress amongst infants in rural India: findings from the early life stress sub-study of the SPRING cluster randomised controlled trial (SPRING-ELS).** *Psychoneuroendocrinology.* 2019; **107**(November 2018): 241–250.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bristow M, Clemens L, Parker K: **Hair cortisol concentration method validation study 2.** *figshare. Dataset.* 2020.
[Publisher Full Text](#)
- Clow A, Thorn L, Evans P: **The awakening cortisol response: methodological issues and significance.** *Stress (Amsterdam, Netherlands).* 2004; **7**(March): 29–37.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Dettenborn L, Tietze A, Bruckner F, *et al.*: **Higher cortisol content in hair among long-term unemployed individuals compared to controls.** *Psychoneuroendocrinology.* 2010; **35**(9): 1404–1409.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Dickerson SS, Kemeny ME: **Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research.** *Psychol Bull.* 2004; **130**(3): 355–391.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Feng C, Wang H, Lu N, *et al.*: **Log-transformation and its implications for data analysis.** *Shanghai Arch Psychiatry.* 2014; **26**(2): 105.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Faul F, Erdfelder E, Lang AG, *et al.*: **G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences.** *Behav Res Methods.* 2007; **39**(2): 175–191.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Gao W, Stalder T, Foley P, *et al.*: **Quantitative analysis of steroid hormones in human hair using a column-switching LC-APCI-MS/MS assay.** *J Chromatogr B Analyt Technol Biomed Life Sci.* 2013.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Grouven U, Bender R, Ziegler A, *et al.*: **Comparing methods of measurement.** *Deutsche Medizinische Wochenschrift.* 1946; **132** Suppl(24), 69–73.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hucklebridge F, Hussain T, Evans P, *et al.*: **The diurnal patterns of the adrenal steroids cortisol and dehydroepiandrosterone (DHEA) in relation to awakening.** *Psychoneuroendocrinology.* 2005; **30**: 51–57.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Manenschijs L, Koper JW, Lamberts SW, *et al.*: **Evaluation of a method to measure long term cortisol levels.** *Steroids.* 2011; **76**(10–11): 1032–1036.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Miller R, Plessow F, Rauh M, *et al.*: **Comparison of salivary cortisol as measured by different immunoassays and tandem mass spectrometry.** *Psychoneuroendocrinology.* 2013; **38**: 50–57.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Parker K, Bristow M: **Hair cortisol analysis protocol.** *protocols.io* 2020.
[Publisher Full Text](#)
- Raul JS, Cirimele V, Ludes B, *et al.*: **Detection of physiological concentrations of cortisol and cortisone in human hair.** *Clin Biochem.* 2004; **37**(12): 1105–1111.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Reid J, Parker K, Bristow M: **Hair cortisol concentration method validation study 1.** *figshare. Dataset.* 2020.
[Publisher Full Text](#)
- Sauvé B, Koren G, Walsh G, *et al.*: **Measurement of cortisol in human hair as a biomarker of systemic exposure.** *Clin Invest Med.* 2007; **30**(5): 183–191.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Skoluda N, Dettenborn L, Stalder T, *et al.*: **Elevated hair cortisol concentrations in endurance athletes.** *Psychoneuroendocrinology.* 2012; **37**(5): 611–617.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Stalder T, Kirschbaum C, Heinze K, *et al.*: **Use of hair cortisol analysis to detect hypercortisolism during active drinking phases in alcohol-dependent individuals.** *Biol Psychol.* 2010; **85**(3): 357–360.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Stalder T, Tietze A, Steudte S, *et al.*: **Elevated hair cortisol levels in chronically stressed dementia caregivers.** *Psychoneuroendocrinology.* 2014; **47**(0): 26–30.
[PubMed Abstract](#) | [Publisher Full Text](#)

Stalder T, Steudte-Schmiedgen S, Alexander N, *et al.*: **Stress-related and basic determinants of hair cortisol in humans: A meta-analysis.** *Psychoneuroendocrinology*. 2017; **77**: 261–274.
[PubMed Abstract](#) | [Publisher Full Text](#)

Stetler CA, Guinn V: **Cumulative cortisol exposure increases during the academic term: Links to performance-related and social-evaluative stressors.** *Psychoneuroendocrinology*. 2020; 114.
[PubMed Abstract](#) | [Publisher Full Text](#)

Steudte S, Kolassa IT, Stalder T, *et al.*: **Increased cortisol concentrations in hair of severely traumatized Ugandan individuals with PTSD.** *Psychoneuroendocrinology*. 2011; **36**(8): 1193–1200.
[PubMed Abstract](#) | [Publisher Full Text](#)

Van Uum SH, Sauve B, Fraser LA, *et al.*: **Elevated content of cortisol in hair of patients with severe chronic pain: a novel biomarker for stress.** *Stress*. 2008; **11**(6): 483–488.
[PubMed Abstract](#) | [Publisher Full Text](#)

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research