

OPEN

A systematic review of studies utilizing hair glucocorticoids as a measure of stress suggests the marker is more appropriate for quantifying short-term stressors

Otto Kallioikoski¹, Finn K. Jellestad² & Robert Murison²

Quantitating glucocorticoids (GCs) in hairs is a popular method for assessing chronic stress in studies of humans and animals alike. The cause-and-effect relationship between stress and elevated GC levels in hairs, sampled weeks later, is however hard to prove. This systematic review evaluated the evidence supporting hair glucocorticoids (hGCs) as a biomarker of stress. Only a relatively small number of controlled studies employing hGC analyses have been published, and the quality of the evidence is compromised by unchecked sources of bias. Subjects exposed to stress mostly demonstrate elevated levels of hGCs, and these concentrations correlate significantly with GC concentrations in serum, saliva and feces. This supports hGCs as a biomarker of stress, but the dataset provided no evidence that hGCs are a marker of stress outside of the immediate past. Only in cases where the stressor persisted at the time of hair sampling could a clear link between stress and hGCs be established.

Measuring glucocorticoids (GCs) deposited in hair is an increasingly popular method for biomarker-based stress assessment. Hair is sampled easily and painlessly, it is often an abundant source of material, and it has been argued to have superior qualities over other methods for analyzing GCs when it comes to gauging chronic stress^{1–3}. If GCs are sequestered from the blood stream and locked into place in the growing hair at the level of the hair follicle, a single strand of hair contains within it a historical record of the HPA axis activity of its owner spanning months into the past. This idea is taken to the next level when researchers segment hairs and analyze different sections, ostensibly corresponding to different periods in the past, to make inferences regarding the perceived stress levels over time of their subjects, whether human patients⁴, captive animals⁵, or long-dead mummies⁶.

But are hair glucocorticoids (hGCs) a robust marker of stress? Local production of GCs in the hair follicle has been confirmed⁷, with the local (follicular) HPA axis appearing to respond to local stressors independently of the rest of the organism⁸. The rate of incorporation of GCs into hairs⁹ is unclear and the mechanisms by which this takes place are unknown^{10,11}. Taken together, it is unclear to which degree hGCs are reflective of the (central) stress response of an individual. It has even been suggested that an individual's hGC levels may follow a circadian rhythm, changing with the time of day¹². Moreover, with some evidence that hGC concentrations change much faster than can be explained by mechanisms concerning only incorporation of GCs in the hair follicle^{13,14}, can sections of a hair really be related to a specific period in the past?

Hair analysis for quantitating GC levels builds on a long-standing forensic practice of looking for deposition of toxins and illicit substances in hairs and fingernails¹⁵. Methods for detecting residual anabolic steroid hormones in hair gained momentum in the wake of a high-profile doping scandal in competitive cycling in 1998^{16,17}, and were subsequently further extended for quantitating also endogenously produced steroids¹⁸. A study in 2002 on rock hyraxes² is, we believe, the first study to infer preceding stress from the concentrations of hGCs. In 2007 (the study seeming to have been initiated years earlier¹⁹), the first study linking hGCs to stress was published for humans²⁰, citing the method's successful application in animals (specifically, Davenport *et al.*¹). Notably, with a

¹Department of Experimental Medicine, University of Copenhagen, Copenhagen, Denmark. ²Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway. Correspondence and requests for materials should be addressed to O.K. (email: ohk@sund.ku.dk)

leg-up with respect to method development from forensics, none of these early papers needed to address methodological issues with respect to when and how GCs were sequestered into hairs. There are, however, crucial differences between the forensic application of hair analysis and hair analysis of GCs for estimating preceding stress. With illicit (exogenous) substances, concentrations are less important – there are no legal circulating levels of cocaine, athletes should not present with traces of nandrolone in their system, and if your hairs contain traces of thallium you have angered a regime and you are secretly being poisoned – quantitation is mostly only relevant with regards to detection limits. Endogenous compounds are more problematic in a forensic context as relative levels become important (testosterone doping in athletes being a notorious forensic conundrum²¹). Moreover, relating the position in a strand of hair of an exogenous substance of interest to a time of exposure in the past is highly approximate, even if the substance's interactions with the hair matrix are well-characterized^{22,23}. Individual differences in hair uptake (capacity) have also been noted, suggesting that comparisons of measured quantities between subjects may be inappropriate^{24,25}. For hGC analyses, by contrast, concentrations, timing of deposition, and inter-individual comparisons are key features. Here, there are no clear precedents within the vast body of literature on forensic hair analysis for hGC analyses to lean on.

Despite the many unknowns surrounding the use of hGCs as a measure of chronic stress, the biomarker is presently used to gauge mental illness²⁶, the wellbeing of human trauma victims²⁷ (and long-term consequences of the trauma^{28,29}), post-traumatic stress disorder (PTSD) sufferers^{4,30,31}, and children^{20,32}; to assess animal welfare in wildlife³³, captive animals^{34,35} and laboratory animals³⁶. The mismatch between the uncertainties of the method and the confidence with which it is applied is concerning.

The present systematic review strove to collect and evaluate the empirical evidence supporting the use of hGC analysis as a method for assessing physiological and psychological stress. With the method being used both for studying animals and humans, and the two bodies of literature feeding off one another, limiting the study to either humans or non-human animals would not have painted a complete picture. Unlike previous reviews/meta-analyses^{3,26,37,38}, we thus set out to collate data from all vertebrate species; to include studies carried out in human and non-human animals alike. Specifically, we investigated whether hGC levels had been found to correlate well with concentrations of GCs in other biological matrices, proven to be reflective of HPA axis activity in the recent past. We collated studies that compared hGC levels in stressed individuals to those of unstressed controls to determine whether an up-regulated HPA axis would produce a predictable increase in hGCs. Through meta-analytic investigation of these two sets of studies, we attempted to answer whether hGC concentrations appear to be reflective of the central HPA axis activity. Specifically, the focus in this review is to assess the evidence supporting the claim that stressful stimuli result in measurable increases in hGC levels. Moreover, by subdividing controlled studies by type and temporality of the stressor, we attempted to determine whether hGC levels are better suited for describing certain types of stress than others.

Material and Methods

The methods listed below were pre-specified in a study protocol accessible online since Jan 13, 2016 (Supplemental materials C).

A broad – inclusive – search strategy was employed in an attempt to find all relevant publications that could provide unambiguous evidence of hGCs being related to the HPA axis-activating stress response of an individual. Studies where hGC levels were used as a measure of stress or where hGC levels were correlated to GC levels in other biological matrices (blood, saliva, urine and feces) were retrieved through multiple databases (MEDLINE, Web of Science, EMBASE, Zoological Record, and PsycINFO). The publication searches were conducted in January 2016. Whereas we have contextualized our findings with more recent examples, no studies published past this date were included in the analyses. Quality assessments were made according to an adapted (nine-item) checklist and basic study information was extracted along with hGC results. Synthesizing data from multiple sources, summary estimates were created separately for correlations with different biological matrices. Similarly, experimental designs deemed fundamentally incompatible were separated out and individual summary estimates were created. Specifically, studies were classified as studying “induced (acute) stress”, “chronic stress”, “observed stress” (where the stressor was inferred by an observer), “self-assessed stress”, “past stress” (where a subject was exposed to a stressful period, which had subsequently ended prior to hair collection) and post-traumatic stress disorder (“PTSD”; which we chose to include because its link with hGCs was receiving great attention at the time this review was scoped). Random effects models were used throughout and differences in experimental subjects and controls were expressed as standardized mean differences. For detailed descriptions of the methods, refer to Supplemental materials A.

Results

A total of 3,518 unique entries were found using the search strategy, of which 468 entries were retained for full text analysis (Fig. 1). A majority of these studies were subsequently excluded due to not meeting the pre-stated inclusion criteria: 28% were excluded due to their exploratory study design – often characterized by the lack of a control group and a clear *a priori* hypothesis; 16% presented no data from a controlled study – these were mostly method papers, reviews, opinion papers, and other narrative journal entries; 26% were not peer-reviewed publications – these were mostly meeting abstracts and theses. Other incompatible study designs, and entries where the full text could not be obtained, made up 17% of the entries retained for full text screenings.

For the entries retained for full text screening – where all texts were verified to concern the use of hGCs – an exponential growth in method adoption is obvious: 2015 saw more publications on hGCs than had been published between 2003 and 2011 in total. Presently, a new publication (counting also non-peer reviewed entries) on hGCs is available online every three days (or less).

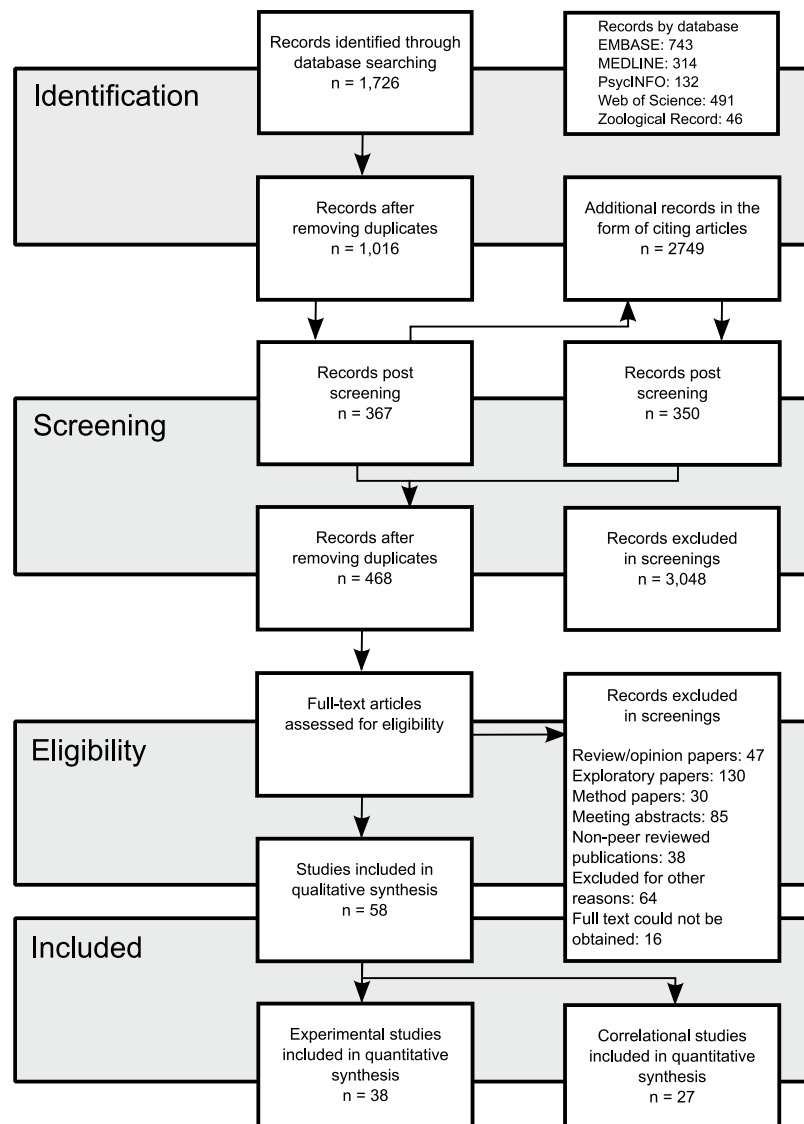


Figure 1. Flow chart outlining the systematic search strategy, the subsequent screening, and inclusion/exclusion of database entries. The diagram has been adapted from the PRISMA Flow Diagram⁸⁷.

Study quality of experimental studies. Of the 59 peer-reviewed publications included in the present systematic review, 38 papers reported on 42 studies with a stress group/control group design that could be assessed for study quality.

A salient trend was found when assessing the risk of bias: A majority of the 38 papers did not account for the possibility that a stressor other than the one that was purportedly studied could have influenced the results. This is evident in Fig. 2 focusing on checklist items 2, 3 and 8: The influence of concurrent interventions or unintended exposures could only be ruled out in 9 (24%) of the studies (item 3), the influence of confounding factors could only be ruled out in 16 (42%) of the studies (item 2), and only 12 (32%) of the studies featured a study design that ensured that the subjects were equally exposed to any confounding factors (item 8). In only three studies (8%) could all three sources of bias be ruled out entirely. Similar ambient conditions for stress and control groups could also only be guaranteed in 15 (39%) of the studies (item 5). Remarkably, only 3 (8%) of the studies reported on blinding of the outcome assessors (item 6), even though this is an explicit recommendation of most present-day best-practice frameworks (e.g. the ARRIVE guidelines³⁹). In no one study were all of the sources of bias addressed, and in a few none were (for a by-entry summary of the risk-of-bias analyses, refer to Supplemental materials B, appendix 1).

Study characteristics and data extraction. The studies retained for analysis presented a diverse set, with no two study designs quite alike (Tables 1 and 2). Of the studies retained for analysis, roughly half (48%) were human studies. Both sexes have been studied in roughly equal numbers (52% female subjects across all studies), but only rarely were equal sex ratios employed in any one study; study objectives and opportunistic sampling of e.g. wildlife populations tending to bias the sex ratio in favor of one or the other. We made initial attempts at

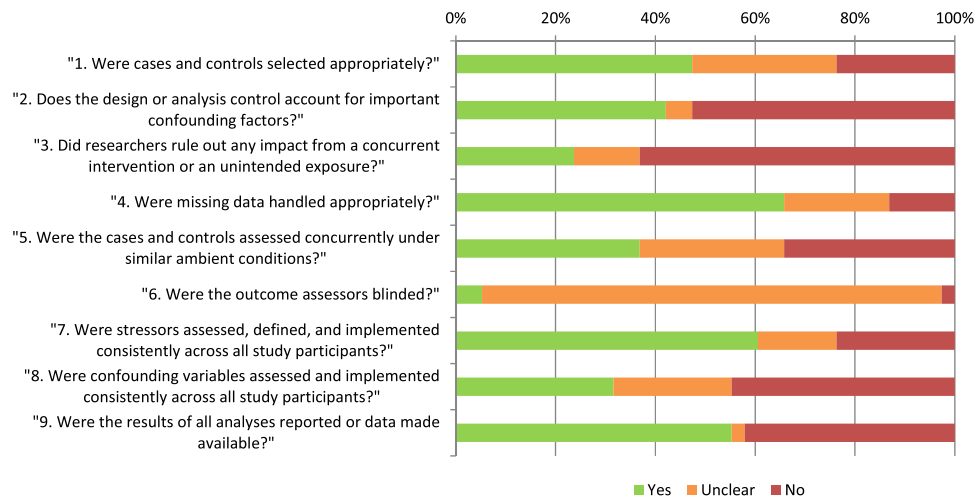


Figure 2. Results from the risk-of-bias checklist assessment of the experimental study designs.

exploring sex differences – similar to a previous meta-analysis³⁸ – however the data were insufficient to draw any conclusions. Similarly, when extracting data we had harbored hopes of being able to compare the effects of differing sampling and analysis protocols that have been discussed previously⁴⁰. However, the laboratory methods employed were fairly similar and study designs fairly dissimilar, the combination lending itself poorly to stringent analyses. Human studies were consistent in sampling the posterior vertex of the head, whereas the non-human studies appeared to sample regions by convenience or just by random (e.g. studies in dogs have sampled backs, shoulders, chests, and legs, depending on research group and study). Although often discussed as a potential issue^{41,42} no one study admitted to including hair follicles in their hair samples and all but six papers^{5,34,43–46} explicitly described methods designed to ensure samples being free of follicles. It has been theorized that the color of a hair can influence the GC content⁴⁷, however the results are inconsistent¹⁰ and hair color is rarely reported. We consequently did not attempt to extract information on this. Only human and other primate studies employed the “stress calendar” idea, sub-sectioning hairs to infer circulating GC levels at multiple time points in the past from the same sample. Of all the studies retained for analysis, a clear majority (48 studies, 83%) employed a washing step, intended to remove contaminants from the outside of the hairs, and all but two studies minced/pulverized the hairs prior to analysis. For quantification of GCs, antibody-based methods were most frequently employed (48 studies, 83%), however numerous different protocols/antibodies have been utilized.

When extracting data, two studies – Manenschiijn *et al.*⁴⁸ and Luo *et al.*⁴ – were singled out as having a reported precision more than tenfold higher than the other 38 studies (including studies utilizing the very same methodology in comparable subjects). We believe that this is simply due to incorrect reporting of the measure of dispersion. Unable to reach the authors for a comment – despite multiple attempts – we have tentatively included the data from these studies, assuming that the graphically presented measures of dispersion were in fact SEMs, rather than – as listed – 95% CIs. Three studies – two experimental studies of chronic stressors^{33,49}, and one study reporting a non-significant correlation between hGC and GC in saliva⁵⁰ – were excluded from further analysis as critical information could not be obtained from the corresponding authors (none of the studies listed the number of samples/subjects used in their analyses). We do not expect these exclusions to have significantly altered our summary estimates, however, as these studies were of moderate size and all fell into well-populated subgroups.

Correlations with GC in other matrices. Meta-analyses of correlation coefficients revealed a great deal of heterogeneity between studies, as could be expected from the diverse set of studies analyzed (Fig. 3). Significant synthesized meta-correlations could be found between hGCs and GCs in blood, saliva and feces. A significant correlation could not be found between hGCs and GCs in urine. However this analysis featured only five studies (collecting 169 subjects), all with fairly high intra-study variance of data. Leave-one-out analysis furthermore revealed that the statistically significant correlation found between GCs in blood and hGCs could not be substantiated if data from the study by Yu *et al.*⁵¹ were removed. Moreover, removing the data from the study by Accorsi *et al.*⁵² would more than halve the synthesized correlation coefficient between GCs in feces and hGC (putting it in range with the other correlations at $r = 0.22$), suggesting that the strength of the correlation may be somewhat overestimated. Additionally, removing a single study for each correlation summary would reduce heterogeneity markedly, bringing down the largest I^2 value to 38%, suggesting a small number of studies were responsible for a majority of the heterogeneity (the leave-one-out analyses are presented in their entirety in Supplemental Materials B, Appendix 2). A factor adding to the heterogeneity may also be that some studies averaged multiple samples over time – e.g. the study by D’Anna-Hernandez *et al.*⁵³ where hair and saliva samples were averaged across four sampling points throughout human pregnancy, or Sauv e *et al.*⁵⁴ comparing single hair samples to urine samples averaged over a 24-hour period. With a relatively small dataset, we did not see fit to analyze these as separate subgroups, increasing our “researcher degrees of freedom”^{55,56}, and potentially flagging spurious correlations

Correlational studies	Study	Subjects	Correlation with GCs in	Males/females/unknown	Sampling site	* = only section nearest to the scalp was analyzed				Analysis method
						Sub-segmentation?	Follicle?	Wash?	Processing?	
Accorsi, 2008	Cats	Feces	8/19/0	Back (ischiatric region)	No	No	No	Cut	RIA	
Accorsi, 2008	Dogs	Feces	21/8/0	Back (ischiatric region)	No	No	No	Cut	RIA	
Bennett, 2010	Dogs	Saliva	0/0/48	Back (ischiatric region)	No	No	Yes	Milled	ELISA	
Bryan, 2013a	Dogs	Feces and saliva	5/2/0	Shoulders	No	No	Yes	Milled	ELISA	
Chan, 2014	Humans	Serum, urine and saliva	26/31/0	Head	No	No	No	Cut	ELISA	
Chen, 2014	Humans	Urine	29/0/0	Head (posterior vertex)	No	No	Yes	Milled	HPLC-MS/MS	
Corradini, 2013	Dogs	Serum	49/41/0	Chest (xiphoid region)	No	No	No	Cut	RIA	
D'Anna-Hernandez, 2011	Humans	Saliva	0/21/0	Head (posterior vertex)	Yes	No	Yes	Milled	ELISA	
Davenport, 2006	Rhesus monkeys	Saliva	20/0/0	Neck (posterior vertex)	Yes	No	Yes	Milled	ELISA	
Kuehl, 2015	Humans	Saliva	31/54/0	Head	Unclear	No	Yes	Unclear	CLIA	
Manenschijn, 2012	Humans	Saliva	0/0/90	Head (posterior vertex)	No	No	No	Cut	ELISA	
Mastromonaco, 2014	Chipmunks	Feces	0/0/62	Leg	No	No	Yes	Cut	ELISA	
Moya, 2013	Cattle	Saliva and feces	12/0/0	Head, neck, shoulder, hip, tail	No	Unclear	Yes	Milled	ELISA	
Moya, 2015	Cattle	Saliva	80/0/0	Tail	No	Unclear	Yes	Milled	ELISA	
Ouschan, 2013	Dogs	Blood	4/8/0	Leg (inside)	No	No	Yes	Cut	ELISA	
Pulopulos, 2014	Humans	Saliva	12/38/0	Head (posterior vertex)	No	No	Yes	Milled	LC-MS/MS	
Sauvé, 2007	Humans	Saliva, urine and serum	19/20/0	Head (posterior vertex)	No	No	No	Cut	ELISA	
Schalinski, 2015	Humans	Saliva	0/28/0	Head (posterior vertex)	Yes	No	Yes	Unclear	CLIA	
Steutde, 2011	Humans	Saliva	0/27/0	Head (posterior vertex)	Yes	No	Yes	Milled	CLIA	
Steutde, 2013	Humans	Saliva	6/72/0	Head (posterior vertex)	Yes	No	Yes	Milled	LC-MS/MS	
Sumra, 2015	Humans	Urine	0/31/0	Head	No*	No	Yes	Ground	ELISA	
Tallo-Parra, 2015	Cattle	Feces	0/17/0	Head	No	No	Yes	Cut	ELISA	
van Holland, 2012	Humans	Saliva	0/0/27	Head	No	No	Yes	Milled	CLIA	
Vanaelst, 2012	Humans	Saliva and serum	0/32/0	Head (posterior vertex)	No*	No	No	Cut	LC-MS/MS	
Wippert, 2014	Humans	Urine	3/10/0	Head (posterior vertex)	No	No	Yes	Unclear	LC-MS/MS	
Xie, 2011	Humans	Saliva	32/0/0	Head (posterior vertex)	No*	No	Yes	Milled	LC-MS/MS	
Yamanashi, 2013	Chimpanzees	Feces	9/0/0	Arm	No	No	Yes	Milled	ELISA	
Yu, 2015	Mice	Serum	19/0/0	Back	No	No	Yes	Cut	LC-MS/MS	
Yu, 2015	Rats	Blood	22/0/0	Back	No	No	Yes	Cut	LC-MS/MS	

Table 1. Study characteristics for studies with extracted correlations. In total, 29 comparisons were extracted from 27 peer-reviewed publications, collecting data from 1,093 subjects across nine species. These studies were used for producing summary estimates of correlation coefficients between hGC concentrations and GC concentrations in other biological matrices.

(this would, further, only have been possible for the correlation with GC in saliva). Moreover, the studies that correlated both point estimates and averages^{31,57–59} did not demonstrate a consistent difference between the two approaches.

hGCs as a measure of stress. Summarizing the evidence from the experimental studies using random effects models produced varied results (Fig. 4), reaffirming our decision to carry out subgroup analyses. Induced (acute) stress models produced a clear elevation in GC concentrations measured in hairs (Fig. 4A) with low inter-study heterogeneity (I^2 could not be estimated). Chronic stressors also produced a significant elevation in deposited GC compared to control groups (Fig. 4B). The results from the chronic stress studies were however highly heterogeneous with a majority of the variance of the summary estimate stemming from between-study variation, as opposed to within-study variation ($I^2 = 80\%$), suggesting that not all of the studies were comparable with respect to the stress response and its effect on hGC levels. Simply put, it is unlikely that these studies all describe a similar HPA axis activation in response to the studied stressor; it is, for instance, likely that some scenarios simply did not induce a stress response. Observed stress (Fig. 4C) and self-assessed stress (Fig. 4D) produced unclear results. Finally, stressors that had subsided at the time of sampling (“past stress”) did not produce a measurable elevation in hGCs (Fig. 4E). Studies concerning hGCs measured in PTSD sufferers similarly generated unclear results (Fig. 4F), with a combination of studies showing both elevations and decreases in hGC output relative to a control group.

Experimental studies	Subjects	Stressor	Classification	Males/ females / unknown	Sampling site	* = only section nearest to the scalp was analyzed				Analysis method
						Sub-segmentation?	Follicle?	Wash?	Processing?	
Ashley, 2011	Caribou	ACTH challenge	Past stress	6/6/0	Neck, shoulder, rump	No	Unclear	Yes	Milled	ELISA
Ashley, 2011	Reindeer	ACTH challenge	Past stress	6/6/0	Neck, shoulder, rump	No	Unclear	Yes	Milled	ELISA
Boesch, 2015	Humans	Military training	Chronic stress	105/0/0	Head (posterior vertex)	No*	No	Yes	Cut	ELISA
Bryan, 2013b	Bears	Anthropogenic disturbance	Chronic stress	87/25/0	Unknown	No	Unclear	Yes	Milled	ELISA
Bryan, 2015	Wolves	Anthropogenic disturbance	Chronic stress	76/72/30	Unknown	No	No	Yes	Milled	ELISA
Carlitz, 2014	Orangutans	Various, undefined	Observed stress	31/27/0	Mixed/unclear	Yes	Unclear	Yes	Cut	CLIA
Cattet, 2014	Bears	Capture stress	Induced (acute) stress	240/246/0	Mixed/unclear	No	No	Yes	Milled	ELISA
Chu, 2014	Cynomolgus monkeys	Post-partum depression	Observed stress	0/10/0	Back	No	No	Yes	Milled	RIA
Corradini, 2013	Dogs	Hypercortisolism	Chronic stress	49/41/0	Chest (xiphoid region)	No	No	No	Cut	RIA
Davenport, 2006	Rhesus monkeys	Relocation stress	Induced (acute) stress	20/0/0	Neck (posterior vertex)	Yes	No	Yes	Milled	ELISA
Dettenborn, 2010	Humans	Long-term unemployment	Chronic stress	13/46/0	Head (posterior vertex)	Yes	No	Yes	Milled	CLIA
Dettmer, 2014	Rhesus monkeys	Crowding stress	Chronic stress	0/0/152	Neck (posterior vertex)	No	No	Yes	Milled	ELISA
Fairbanks, 2011	Vervet monkeys	High stress environment	Induced (acute) stress	0/226/0	Back	No	No	Yes	Milled	ELISA
Fourie, 2015	Vervet monkeys	Anthropogenic disturbance	Chronic stress	29/43/0	Back	No	No	Yes	Cut	ELISA
Gao, 2014	Humans	Traumatic event	Self-assessed stress	75/17/0	Head (posterior vertex)	No	No	Yes	Milled	LC-MS/MS
González-de-la-Vara 2011	Cattle	ACTH challenge	Induced (acute) stress	0/15/0	Unclear	No	No	Yes	Cut	RIA
Heinze, 2016	Humans	Mental health problems	Self-assessed stress	6/52/0	Head (posterior vertex)	Yes	No	Yes	None	CLIA
Henley, 2013	Humans	Socioeconomic stress	Chronic stress	25/15/32	Head (posterior vertex)	No	No	Yes	Cut	ELISA
Jarcho, 2016	Mice	Social instability stress	Induced (acute) stress	0/24/0	Back (lower back)	No	No	Yes	Unclear	ELISA
Kapoor, 2016	Rhesus monkeys	Acoustic startle stress	Past stress	0/35/0	Back (upper back)	No	No	Yes	Milled	LC-MS/MS
Karlén, 2011	Humans	“Serious life-event”	Self-assessed stress	24/71/0	Head (posterior vertex)	No*	No	No	Milled	RIA
Klumbies, 2014	Humans	Social phobia	Observed stress	0/0/53	Head (posterior vertex)	No*	No	Yes	Unclear	CLIA
Luo, 2012	Humans	Traumatic event	Observed stress	0/84/0	Head (posterior vertex)	Yes	No	Yes	Milled	CLIA
Luo, 2012	Humans	Post-traumatic stress disorder	PTSD		Head (posterior vertex)	Yes	No	Yes	Milled	CLIA
Manenschijn, 2011	Humans	Shift work	Chronic stress	122/0/0	Head (posterior vertex)	No	No	No	Unclear	ELISA
Mastromonaco, 2014	Chipmunks	ACTH challenge	Induced (acute) stress	0/0/12	Leg	No	No	Yes	Cut	ELISA
Moya, 2015	Cattle	Digestive problems	Induced (acute) stress	80/0/0	Tail	No	Unclear	Yes	Milled	ELISA
Nejad, 2014	Sheep	Water restriction	Induced (acute) stress	0/9/0	Neck (posterior vertex)	No	No	Yes	Cut	ELISA
Oullette, 2015	Humans	Psychosocial stress	Self-assessed stress	0/60/0	Head (posterior vertex)	No*	No	Yes	Cut	ELISA
Qin, 2015	Rhesus monkeys	Induced SADS	Chronic stress	0/8/0	Back	No	No	Yes	Milled	RIA
Schalinski, 2015	Humans	“Stress-related disorders”	PTSD	0/51/0	Head (posterior vertex)	Yes	No	Yes	Unclear	CLIA
Scorrano, 2015	Rats	Misc. stress protocols	Induced (acute) stress	58/0/0	Back (lower back)	No	No	Yes	Ground	RIA
Skoluda, 2012	Humans	Intensive training	Chronic stress	144/151/0	Head (posterior vertex)	Yes	No	Yes	Milled	CLIA

Continued

Experimental studies	Subjects	Stressor	Classification	Males/ females / unknown	Sampling site	* = only section nearest to the scalp was analyzed				Analysis method
						Sub-segmentation?	Follicle?	Wash?	Processing?	
Stalder, 2014	Humans	Caring for relative with dementia	Chronic stress	4/36/0	Head (posterior vertex)	No*	No	Yes	Milled	CLIA
Steutde, 2013	Humans	Traumatic event	Self-assessed stress	6/72/0	Head (posterior vertex)	Yes	No	Yes	Milled	LC-MS/MS
Steutde, 2013	Humans	Post-traumatic stress disorder	PTSD		Head (posterior vertex)	Yes	No	Yes	Milled	LC-MS/MS
Steutde-Schmiedgen, 2015	Humans	Military deployment	PTSD	90/0/0	Head (posterior vertex)	No*	No	Yes	None	LC-MS/MS
Terwissen, 2013	Lynxes	ACTH challenge	Induced (acute) stress	1/2/0	Unclear	No	Unclear	Yes	Cut	ELISA
van Uum, 2008	Humans	Chronic pain	Chronic stress	25/29/0	Head (posterior vertex)	No*	No	No	Cut	ELISA
Yamada, 2007	Humans	Post-birth complications in neonates	Chronic stress	0/0/78	Head	No	No	No	Cut	ELISA
Yu, 2015	Mice	Aggression/social instability	Induced (acute) stress	19/0/0	Back	No	No	Yes	Cut	LC-MS/MS
Yu, 2015	Rats	Surgery/post-surgical pain	Induced (acute) stress	22/0/0	Back	No	No	Yes	Cut	LC-MS/MS

Table 2. Study characteristics for experimental studies. In total, 42 comparisons were extracted from 38 peer-reviewed publications, collecting data from 3,199 subjects across 16 species. These studies were used for creating summary estimates for the difference in hGC concentrations between stressed subjects and controls for different sub-categories of stressors.

Discussion

Study landscape as evidenced by systematic searches. With a couple of new papers appearing every week that concern or utilize hGC analysis, it is fair to say that it has become a widespread method for assessing stress. But hair-growth is a slow process, and popular speculation^{60,61} suggests that GCs are sequestered by hairs over several weeks – if not months. Consequently, controlled studies are for logistic reasons hard to design and execute. Perhaps this is why our search strategy turned up more narrative reviews, opinion papers, and book chapters lauding the method than it did actual controlled studies providing empirical evidence that the method is a sound one. Moreover, the typical (i.e. the most numerous) study employing hGC analyses, published prior to February 2016, was an exploratory one. Characteristically, a single cohort of subjects had hair samples collected along with a number of other environmental, physiological, psychological, and/or demographic data. Correlations were then constructed to scrutinize which parameters were linked to elevated hGC concentrations. The topics of these studies are varied, from investigations of environmental effects on squirrel gliders⁶² or social effects on German shepherds⁶³ to probing cultural⁶⁴, environmental⁶⁵, nutritional³² or genetic⁶⁶ influences on psychological stress in people of differing ages. The implicit prior assumption for studies of this kind is that hGCs are linked to central HPA axis functioning and are thus a measure of (chronic) stress. This puts even more of an onus on the (relatively small number of) controlled studies to validate and affirm the use of hGC concentrations as a measure of stress.

Is there support for hGCs as a biomarker of stress? The present investigation supports the use of hGCs as a measure of central HPA axis functioning and, consequently, as a stress-sensitive biomarker. The compounded data however calls into question the temporality of the marker, suggesting it is a better marker for ongoing than of past stress.

In studies where subjects were exposed to a controlled stressor, a predictable elevation was found in most cases. Whether repeated ACTH challenges⁶⁷ or a more elaborate protocol combining multiple stressors were employed⁶⁸, a consistent increase was found across species when comparing challenged subjects to unstressed controls. The effect of acute stress on hGC levels seems, furthermore, to be rapid. Whereas most studies sampled hairs at least two weeks after having applied the stressor, the study by Cattet *et al.*¹⁴ is remarkable in that they report elevated hGCs within hours of stressor onset. In a similar vein, most stress protocols were applied continuously for weeks before hGC concentrations were evaluated, but González-de-la-Vara *et al.*⁶⁷ found that they could detect an elevation in hGCs two weeks after a pair of sustained-release ACTH injections. Both studies point to hGC concentrations being reflective, primarily, of events in the recent past, as opposed to historical stressors. This is also consistent with hGCs correlating with GCs in other matrices.

Although both inter- and intra-study variances were high for the collated data, it is clear that hGC concentrations correlate significantly with GC concentrations in other matrices. The synthesized correlation coefficients are weak to moderate – ranging from 0.13 to 0.56 – but this is in range with the correlations between established matrices obtained in these very same studies^{54,59,69,70}. Due to large fluctuations stemming from the pulsatile nature of GC release⁷¹, coupled with the different temporality of the matrices – serum and saliva concentrations of GCs

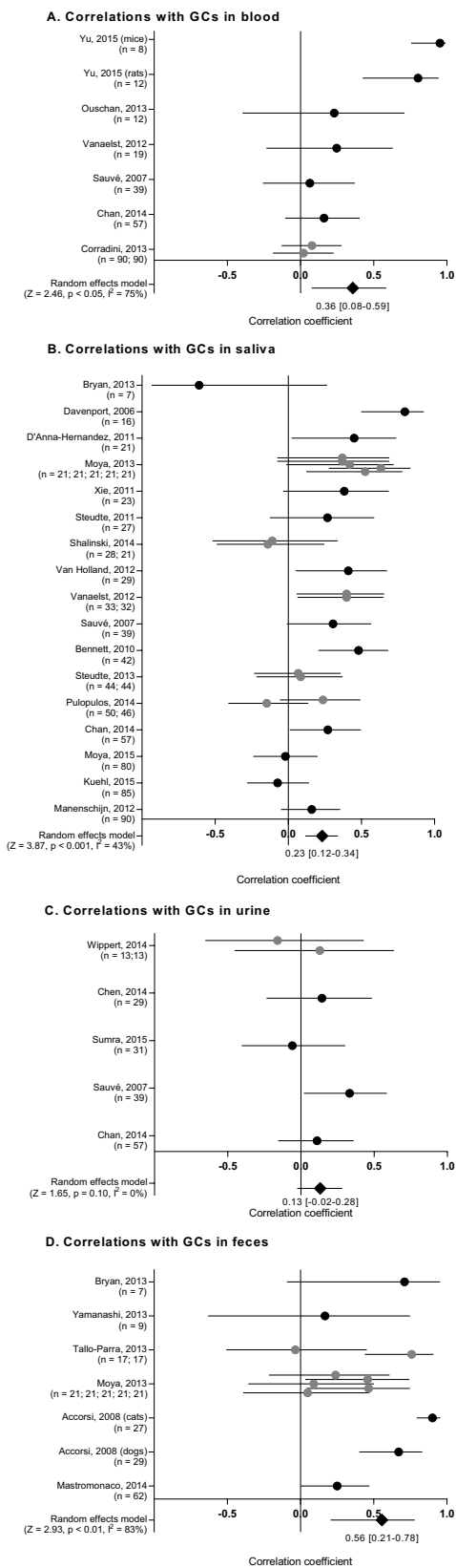


Figure 3. Synthesis of correlation coefficients. Forest plots are presented for correlation coefficients between hGCs and GC in (A) blood, (B) saliva, (C) urine, and (D) feces. Where multiple coefficients were reported in the same study (grey markers) these were weighted accordingly, to avoid biasing the random effects model. The summary estimate correlation coefficients are reported with 95% CI.

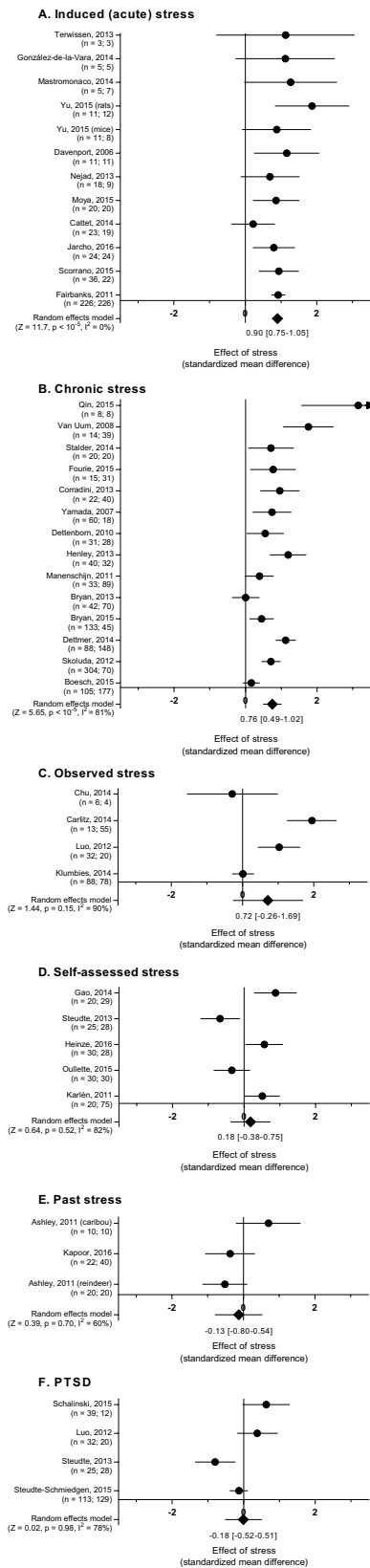


Figure 4. Forest plots summarizing results from induced (acute stress) studies (A), chronic stress studies (B), studies where stress was inferred through observation (C) or assessed by the subject (D), studies where the stressor had passed (E), and studies featuring subjects with PTSD (F). The number of subjects in the studies are listed with the control group last. Summary estimates (Hedges' g) were constructed using the DerSimonian-Laird approach and are reported with 95% CI.

change in a matter of minutes in response to a stressor, urinary and fecal GCs change over a period of hours⁷² – these correlations will inevitably be moderate at the most. The correlation between hGCs and GCs in feces is the strongest of the four, which is to be expected as fecal samples integrate circulating GC concentrations over a period of several hours. Hairs are similarly suggested to sequester GCs from circulation over a longer time window. In the face of popular claims, it is unlikely that this time window is several weeks long, however, as hGC concentrations also correlate significantly with serum and salivary concentrations of GCs.

The effect of confounders on measured hGC levels in chronic stress studies. When compiling studies of chronic stress, a link between individuals experiencing stress and elevated levels of hGCs was found, albeit a weaker link than for acute stressors. The greater level of heterogeneity of this dataset is probably in part because some of the studies were carried out under highly uncontrolled circumstances. With long-term studies featuring subjects – whether human or non-human – in an uncontrolled environment, it is hard to ensure that the studied stressor is the sole and most influential source of stress. It may be that a lack of dietary salmon elicits a physiological stress reaction in grizzly bears, as suggested by Bryan *et al.*⁴³, but it is quite impossible to tell what other factors might influence the life and allostasis of these bears. The confounding factors of this study may well have overshadowed the effect the authors were looking for. Similarly, military training is not all long marches and adrenaline-fueled combat training. With no outside verification, the soldiers undergoing basic training studied by Boesch and collaborators⁷³ may not have had a more active HPA axis than e.g. an office worker with an active lifestyle in the period of sampling. This is not to criticize these experiments; rather, this is to highlight the fact that a number of studies into chronic stressors have an exploratory element to them, as the magnitude of the chronic stressor is hard to judge in relation to a host of ambient stressors. Our risk-of-bias assessment singled out unrelated confounding factors as the most common unchecked source of bias. Only 24% of studies could account for external confounding factors in the studied period, and only in 32% of studies could they be assumed to have been distributed equally between the studied subject groups. The evidence supplied by the chronic stress studies should thus be interpreted carefully.

Differences between the human concept of stress and HPA axis activity. In our investigation, studies where periods of stress were inferred presented highly heterogeneous data. It has been shown before that when human subjects are asked to introspectively assess their own level of stress, assessments correlate poorly with their actual HPA axis functioning^{38,74,75}. In the present investigation we see a similar trend for studies relying on self-reported stress. Whereas we will note that the present investigation contains only a handful of studies, no consistent trend or even weak effect can be inferred. This is not to say that the subjects were not experiencing psychological stress – the studies collect data from distressed subjects ranging from survivors of natural disasters⁷⁶ to patients sourced from mental health services^{31,77} – but it serves as a reminder that the human concept of stress is not synonymous with the prototypical fight-or-flight response. Different states of stress will involve the HPA axis differently. This is further exemplified by the studies of PTSD, where in two studies^{31,78} a notable reduction in hGCs is found for PTSD subjects when compared to healthy controls. We specifically analyzed PTSD studies separately as it has been suggested that PTSD is accompanied by a lowering in circulating GC levels, as opposed to an elevation. Notably, PTSD subjects are identified through clinical scores suggesting the subjects were assigned to groups according to arbitrary cutoffs in a continuum of chronic stress diagnoses. This muddying of the waters, where the line between chronic stress conditions and PTSD is blurred, may, in part, explain why no clear trend is found concerning hGC profiles for either. In the future, a larger dataset that would allow for a more stringent subgrouping of chronic stress studies based on e.g. clinical scores may assist in identifying more uniform profiles.

Studies where the stressed subjects were identified through observation similarly did not paint a consistent picture. Whether the result of studying animal behavior^{5,79} or of putting human patients through structured interviews^{4,80}, there seems to be a mismatch between subjectively assessed stress and hGC levels. For this category of studies we will note that it is particularly concerning that no blinding was employed, even though the findings hinge completely on the subjective assessment of an external observer. Would the marked difference between groups have been as profound in the study by Carlitz *et al.*⁵ (the average effect size was greater than that found in any induced stress study) if periods of stress had been recorded by a blinded observer? The study was made even more subjective and bias-prone by not ever defining the studied stressors, leaving it in the hands of unblinded observers to determine what to consider a stressor. With few studies and heterogeneous results, it is currently hard to determine whether studies that rely on externally assessing stress can provide empirical evidence with respect to the utility of hGC analyses.

Temporal aspects of hGCs as a biomarker of stress. An important factor shared between the chronic stress studies that demonstrate a clear difference between stressed and control subjects is that the stressor persisted at the time of sampling. When singling out the studies where the stressor could be positively ensured to have subsided at the time of hair sampling, the pattern was found to be different. In the study by Kapoor *et al.*⁸¹, pregnant rhesus monkeys were exposed to a daily acoustic startle stress protocol for five weeks. Serum samples analyzed for circulating GC levels were used to verify that the protocol elicited a significant stress response throughout the period. When analyzing hGC concentrations 3–13 weeks later (depending on subject), no elevation could be found relative to a control group; not a trace to be found of a considerable elevation of circulating GC levels persisting for five weeks. Similarly, when Ashley *et al.*³⁴ analyzed hairs from both reindeer and caribou two weeks after a single (non-sustained-release) ACTH challenge, no elevation could be found. Fecal GC analyses confirmed that the stressor had subsided after 24–48 hours. With only two studies in this category, we should be careful not to over-interpret; however, this is all part of a recurring pattern. In a recent meta-analysis, Stalder *et al.*³⁸ reanalyzed historical data from human studies in aggregate – collecting data from 66 studies and more than 10,000 hGC samples – and found that in cases of past/absent stress, no relation with elevated hGC concentrations

could be found. The idea of hairs containing a historical record of past stress is, and remains, completely unproven, empirical evidence instead pointing to hGCs being a measure of concurrent stress.

Evidence that hGC concentrations are a historical record of stress could also come in the form of studies sub-sectioning hairs, inferring circulating GC levels at multiple time-points in the past. However, the GC levels of hairs were found to be similar across all the sampled segments – individuals with elevated levels of hGCs had higher levels of hGCs in all segments when compared to controls^{77,82,83}. In only two studies the authors attempted to construct a narrative based on point-to-point fluctuations in GC concentrations along the hair shafts. The findings by Luo and collaborators⁴ are however marred by a strong wash-out effect, with hGC levels successively becoming lower the further away from the scalp a segment is sourced. The most distant segment is purported to contain the lowest levels of hGCs as this segment is hypothesized to correspond to a period before a major trauma. However, this also holds true for the non-traumatized controls, undermining the hypothesis. The study by Carlitz *et al.*⁵ is similarly problematic in that the narrative seems to have been constructed *post hoc*, and only three individual profiles are shown in the paper. To our knowledge, there is little evidence to suggest that interactions between a steroid hormone and a strand of hair are strong enough to lock the molecules permanently in place. This effect is the basis by which a specific section of hair can be related to a time-point in the past. Convincing evidence that baleen from whales can trap hormones, leaving a historical record of hormonal fluctuations, has been presented^{84,85}. A similar case for hair – a distantly related keratinous matrix – remains elusive however. A recent investigation using radiolabeled GCs in monkeys has instead presented fairly conclusive evidence that GCs do not form discrete bands in strands of hair, but that GCs move along the shaft of hair post-deposition⁸⁶. The difference may lie in the gauge and density of the matrices, with baleen samples being extracted from a depth of several centimeters, using power tools, as opposed to processing the entirety of a micrometer-thick hair.

Regardless, the evidence provided by sub-sectioning of hairs, taken altogether, rather seems to suggest that hGCs are distributed along a strand of hair by longitudinal transport of the hydrophobic hormones, whether through diffusion or capillary action (possibly helped along by the waxy sebum). Reading too much into point-to-point fluctuations thus currently appears to be a case of seeing patterns where there are none to be found.

Conclusions

Combining results of controlled studies with the correlational evidence, it seems fair to state that hGC levels seem to relate to central HPA axis functioning. GC levels in hairs appear to be an appropriate marker of *ongoing* physiological stress. If the stressor persists, hGC analyses will remain useful; however, it is currently unadvisable to interpret events in the past based on hGC levels. The idea of GCs being locked into place, providing a historical record of HPA axis functioning has been called into question every time it has been tested in a controlled experiment. Based on the collected evidence we would strongly advise against sub-segmenting hairs, speculating about specific periods in the past. We would be delighted to be proven wrong by a future study, but there is something to be said about, not only the studies our search strategy uncovered, but also the ones that could not be found. Whereas it is hard to design a study where subjects' stress levels are controlled for weeks on end, it is far from impossible to design a study to test the hypothesis that a stressor in the past can be uncovered in a specific segment of hair. Yet, these studies are nowhere to be found. Whereas the data material did not allow for a stringent exploration of publication bias, it seems highly probable that a number of studies providing negative results have remained unpublished. With this review, and others like it, it is our hope that these negative findings may find their way into publications, providing a better picture of when hGC analyses are appropriate, and when they are not.

We strongly recommend that current and future research into hGC analyses focus on some of the fundamental questions. How are GCs incorporated into hairs? How long do they remain post-deposition? There are a number of basic questions that can be answered by small means, utilizing clever study design. With two notable exceptions^{9,86}, studies utilizing radioisotope-labelled GCs are virtually completely missing; yet, the information that could be gained from studies of this type is invaluable. Applied studies utilizing hGC analyses, in the meanwhile, would do well to approach theoretical concepts surrounding hGCs in an agnostic fashion. We do not currently know how far into the past we can measure stress through sampling hairs. Stating that a certain protocol measures e.g. three months of preceding stress is misleading and perpetuates misinformation. Moreover, it is important that the duration of stressors be recorded and reported as accurately as possible. If we are able to pin down the temporality of hGCs as a stress marker, findings of studies in the past may have to be re-interpreted. This will however only be possible if there is enough information on timing of stressors relative to hair samplings. By detailing, as best as they can, the timing of stressors, researchers will, in a sense, future-proof their study results. Moreover, we hope that researchers will be more wary of unaccounted-for sources of stress in their studies. In many cases, these cannot be avoided. Consequently, we would encourage the reporting of possible “contaminating factors” – sources of stress that could not be accounted for in the study design. We would also urge authors to publish their raw data, or, at the very least, keep a record of them. In our investigation we were disappointed to find peer-reviewed publications missing crucial basic information (e.g. the number of subjects analyzed in a study) and to learn, when contacting the authors, that the information could not be produced. With most journals being able to host supporting data files online, and a host of repositories available for when journals cannot, there is no reason for not making data available and risking losing important records. Finally, we will note that many of the analyzed papers utilizing hGC analyses have made useful contributions to science; whether giving a voice to overlooked wildlife, trying to improve animal welfare, or assessing the mental wellbeing of people. We may seem critical of some of these studies, however, this comes from an adamant belief that we can and should do even better. We must constantly hold ourselves to a higher standard, in order to improve our field of research.

References

- Davenport, M. D., Tiefenbacher, S., Lutz, C. K., Novak, M. A. & Meyer, J. S. Analysis of endogenous cortisol concentrations in the hair of rhesus macaques. *Gen. Comp. Endocrinol.* **147**, 255–261 (2006).
- Koren, L. *et al.* A novel method using hair for determining hormonal levels in wildlife. *Anim. Behav.* **63**, 403–406 (2002).
- Wosu, A. C., Valdimarsdóttir, U., Shields, A. E., Williams, D. R. & Williams, M. A. Correlates of cortisol in human hair: implications for epidemiologic studies on health effects of chronic stress. *Ann. Epidemiol.* **23**, 797–811.e792 (2013).
- Luo, H. *et al.* Hair cortisol level as a biomarker for altered hypothalamic-pituitary-adrenal activity in female adolescents with posttraumatic stress disorder after the 2008 Wenchuan earthquake. *Biological Psychiatry* **72**, 65–69, <https://doi.org/10.1016/j.biopsych.2011.12.020> (2012).
- Carlitz, E. H., Kirschbaum, C., Stalder, T. & van Schaik, C. P. Hair as a long-term retrospective cortisol calendar in orang-utans (*Pongo spp.*): New perspectives for stress monitoring in captive management and conservation. *Gen. Comp. Endocrinol.* **195**, 151–156 (2014).
- Loerbroks, A., Hoffmann, F., Grimm, A. & Kirschbaum, C. Stressful ancient Egypt? Assessing cortisol concentrations in a Mummy's hair. *Psychosomatic Medicine* **73**(3), A89 (2011).
- Ito, N. *et al.* Human hair follicles display a functional equivalent of the hypothalamic-pituitary-adrenal axis and synthesize cortisol. *The FASEB journal* **19**, 1332–1334 (2005).
- Salaberger, T. *et al.* Influence of external factors on hair cortisol concentrations. *Gen. Comp. Endocrinol.* **233**, 73–78 (2016).
- Keckeis, K. *et al.* Hair cortisol: a parameter of chronic stress? Insights from a radiometabolism study in guinea pigs. *Journal of Comparative Physiology B* **182**, 985–996 (2012).
- Heimbürge, S., Kanitz, E. & Otten, W. The use of hair cortisol for the assessment of stress in animals. *Gen. Comp. Endocrinol.* **270**, 10–17 (2019).
- Sharpley, C. F., McFarlane, J. R. & Slominski, A. Stress-linked cortisol concentrations in hair: what we know and what we need to know. *Reviews in the Neurosciences* **23**, 111–121, <https://doi.org/10.1515/rns.2011.058> (2012).
- Sharpley, C. F., Kauter, K. G. & McFarlane, J. R. Diurnal variation in peripheral (hair) vs central (saliva) HPA axis cortisol concentrations. Clinical medicine insights. *Endocrinology and diabetes* **3**, 9 (2010).
- Sharpley, C. F., Kauter, K. G. & McFarlane, J. R. An Initial Exploration of *in vivo* Hair Cortisol Responses to a Brief Pain Stressor: Latency, Localization and Independence Effects. *Physiological Research* **58**, 757–761 (2009).
- Cattet, M. *et al.* Quantifying long-term stress in brown bears with the hair cortisol concentration: a biomarker that may be confounded by rapid changes in response to capture and handling. *Conservation Physiology* **2**, cou026 (2014).
- Sachs, H. History of hair analysis. *Forensic science international* **84**, 7–16 (1997).
- Gaillard, Y., Vayssette, F. & Pépin, G. Compared interest between hair analysis and urinalysis in doping controls: Results for amphetamines, corticosteroids and anabolic steroids in racing cyclists. *Forensic science international* **107**, 361–379 (2000).
- Stokes, S. In *Velonation* (2013).
- Cirimele, V., Kintz, P., Dumestre, V., Goulle, J. & Ludes, B. Identification of ten corticosteroids in human hair by liquid chromatography–ionspray mass spectrometry. *Forensic science international* **107**, 381–388 (2000).
- Yamada, J., Stevens, B., de Silva, N., Klein, J. & Koren, G. Hair cortisol as a biologic marker of chronic stress in neonates: A pilot study. *Pediatric Research* **53**, 454A–454A (2003).
- Yamada, J. *et al.* Hair cortisol as a potential biologic marker of chronic stress in hospitalized neonates. *Neonatology* **92**, 42–49, <https://doi.org/10.1159/000100085> (2007).
- Kicman, A. T. & Cowan, D. A. Subject-based profiling for the detection of testosterone administration in sport. *Drug Test. Anal.* **1**, 22–24 (2009).
- LeBeau, M. A., Montgomery, M. A. & Brewer, J. D. The role of variations in growth rate and sample collection on interpreting results of segmental analyses of hair. *Forensic science international* **210**, 110–116 (2011).
- Sachs, H. Theoretical limits of the evaluation of drug concentrations in hair due to irregular hair growth. *Forensic science international* **70**, 53–61 (1995).
- Mieczkowski, T. & Newel, R. Statistical examination of hair color as a potential biasing factor in hair analysis. *Forensic science international* **107**, 13–38 (2000).
- Wennig, R. Potential problems with the interpretation of hair analysis results. *Forensic science international* **107**, 5–12 (2000).
- Staufenbiel, S. M., Penninx, B. W., Spijker, A. T., Elzinga, B. M. & van Rossum, E. F. Hair cortisol, stress exposure, and mental health in humans: a systematic review. *Psychoneuroendocrinology* **38**, 1220–1235 (2013).
- Karlén, J., Ludvigsson, J., Frostell, A., Theodorsson, E. & Faresjö, T. Cortisol in hair measured in young adults—a biomarker of major life stressors? *BMC clinical pathology* **11**, 12 (2011).
- Hinkelmann, K. *et al.* Association between childhood trauma and low hair cortisol in depressed patients and healthy control subjects. *Biol Psychiatry* **74**, e15–e17 (2013).
- Grassi-Oliveira, R. *et al.* Hair cortisol and stressful life events retrospective assessment in crack cocaine users. *The American journal of drug and alcohol abuse* **38**, 535–538 (2012).
- Steutde, S. *et al.* Increased cortisol concentrations in hair of severely traumatized Ugandan individuals with PTSD. *Psychoneuroendocrinology* **36**, 1193–1200 (2011).
- Steutde, S. *et al.* Hair Cortisol as a Biomarker of Traumatization in Healthy Individuals and Posttraumatic Stress Disorder Patients. *Biological Psychiatry* **74**, 639–646, <https://doi.org/10.1016/j.biopsych.2013.03.011> (2013).
- Vaghri, Z. *et al.* Hair cortisol reflects socio-economic factors and hair zinc in preschoolers. *Psychoneuroendocrinology* **38**, 331–340 (2013).
- Mastromonaco, G. F., Gunn, K., McCurdy-Adams, H., Edwards, D. & Schulte-Hostedde, A. I. Validation and use of hair cortisol as a measure of chronic stress in eastern chipmunks (*Tamias striatus*). *Conservation Physiology* **2**, cou055 (2014).
- Ashley, N. *et al.* Glucocorticosteroid concentrations in feces and hair of captive caribou and reindeer following adrenocorticotrophic hormone challenge. *Gen. Comp. Endocrinol.* **172**, 382–391 (2011).
- Detmer, A. M., Novak, M. A., Suomi, S. J. & Meyer, J. S. Physiological and behavioral adaptation to relocation stress in differentially reared rhesus monkeys: Hair cortisol as a biomarker for anxiety-related responses. *Psychoneuroendocrinology* **37**, 191–199 (2012).
- Scorrano, F. *et al.* Validation of the long-term assessment of hypothalamic–pituitary–adrenal activity in rats using hair corticosterone as a biomarker. *The FASEB Journal*, fj. 14-254474 (2014).
- Herane Vives, A. *et al.* The relationship between cortisol, stress and psychiatric illness: New insights using hair analysis. *Journal of Psychiatric Research* **70**, 38–49, <https://doi.org/10.1016/j.psychires.2015.08.007> (2015).
- Stalder, T. *et al.* Stress-related and basic determinants of hair cortisol in humans: a meta-analysis. *Psychoneuroendocrinology* **77**, 261–274 (2017).
- Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol.* **8**, e1000412 (2010).
- Russell, E. *et al.* Toward Standardization of Hair Cortisol Measurement: Results of the First International Interlaboratory Round Robin. *Therapeutic Drug Monitoring* **37**, 71–75 (2015).
- Gow, R., Thomson, S., Rieder, M., Van Uum, S. & Koren, G. An assessment of cortisol analysis in hair and its clinical applications. *Forensic science international* **196**, 32–37 (2010).

42. Cook, N. J. Review: Minimally invasive sampling media and the measurement of corticosteroids as biomarkers of stress in animals. *Canadian Journal of Animal Science* **92**, 227–259, <https://doi.org/10.4141/cjas2012-045> (2012).
43. Bryan, H. M., Darimont, C. T., Paquet, P. C., Wynne-Edwards, K. E. & Smits, J. E. G. Stress and Reproductive Hormones in Grizzly Bears Reflect Nutritional Benefits and Social Consequences of a Salmon Foraging Niche. *PLoS ONE* **8**(e80537), 80531–80510 (2013).
44. Moya, D., Schwartzkopf-Genswein, K. S. & Veira, D. M. Standardization of a non-invasive methodology to measure cortisol in hair of beef cattle. *Livestock Science* **158**, 138–144, <https://doi.org/10.1016/j.livsci.2013.10.007> (2013).
45. Moya, D. *et al.* Effect of grain type and processing index on growth performance, carcass quality, feeding behavior, and stress response of feedlot steers. *Journal of Animal Science* **93**, 3091–3100, <https://doi.org/10.2527/jas2014-8680> (2015).
46. Terwissen, C. V., Mastromonaco, G. F. & Murray, D. L. Influence of adrenocorticotrophin hormone challenge and external factors (age, sex, and body region) on hair cortisol concentration in Canada lynx (*Lynx canadensis*). *General and Comparative Endocrinology* **194**, 162–167 (2013).
47. Neumann, A. *et al.* Predicting hair cortisol levels with hair pigmentation genes: a possible hair pigmentation bias. *Scientific reports* **7**, 8529 (2017).
48. Manenschijn, L., Van Kruysbergen, R. G. P. M., De Jong, F. H., Koper, J. W. & Van Rossum, E. F. C. Shift work at young age is associated with elevated long-term cortisol levels and body mass index. *Journal of Clinical Endocrinology and Metabolism* **96**, E1862–E1865, <https://doi.org/10.1210/jc.2011-1551> (2011).
49. Qin, D. *et al.* Cortisol responses to chronic stress in adult macaques: moderation by a polymorphism in the serotonin transporter gene. *Behav Brain Res* **278**, 280–285, <https://doi.org/10.1016/j.bbr.2014.10.001> (2015).
50. Kamps, A. W. A. *et al.* Children with asthma have significantly lower long-term cortisol levels in their scalp hair than healthy children. *Acta Paediatrica* **103**, 957–961, <https://doi.org/10.1111/apa.12685> (2014).
51. Yu, T. *et al.* Determination of endogenous corticosterone in rodent's blood, brain and hair with LC-APCI-MS/MS. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **1002**, 267–276, <https://doi.org/10.1016/j.jchromb.2015.08.035> (2015).
52. Accorsi, P. A. *et al.* Cortisol determination in hair and faeces from domestic cats and dogs. *General and Comparative Endocrinology* **155**, 398–402, <https://doi.org/10.1016/j.ygcen.2007.07.002> (2008).
53. D'Anna-Hernandez, K. L., Ross, R. G., Natvig, C. L. & Laudenslager, M. L. Hair cortisol levels as a retrospective marker of hypothalamic–pituitary axis activity throughout pregnancy: comparison to salivary cortisol. *Physiology & behavior* **104**, 348–353 (2011).
54. Sauve, B., Koren, G., Walsh, G., Tokmakejian, S. & Van Uum, S. H. M. Measurement of cortisol in human hair as a biomarker of systemic exposure. *Clin Invest Med* **30**, E183–E191 (2007).
55. Simmons, J. P., Nelson, L. D. & Simonsohn, U. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* **22**, 1359–1366 (2011).
56. Gelman, A. & Loken, E. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. Department of Statistics, Columbia University (2013).
57. Pulpulos, M. M. *et al.* Hair cortisol and cognitive performance in healthy older people. *Psychoneuroendocrinology* **44**, 100–111 (2014).
58. Schalinski, I., Elbert, T., Steudte-Schmiedgen, S. & Kirschbaum, C. The Cortisol Paradox of Trauma-Related Disorders: Lower Phasic Responses but Higher Tonic Levels of Cortisol Are Associated with Sexual Abuse in Childhood. *Plos One* **10**, <https://doi.org/10.1371/journal.pone.0136921> (2015).
59. Vanaelst, B. *et al.* Intercorrelations between serum, salivary, and hair cortisol and child-reported estimates of stress in elementary school girls. *Psychophysiology* **49**, 1072–1081 (2012).
60. Russell, E., Koren, G., Rieder, M. & Van Uum, S. Hair cortisol as a biological marker of chronic stress: current status, future directions and unanswered questions. *Psychoneuroendocrinology* **37**, 589–601 (2012).
61. Stalder, T. & Kirschbaum, C. Analysis of cortisol in hair—State of the art and future directions. *Brain, behavior, and immunity* **26**, 1019–1029 (2012).
62. Brearley, G., McAlpine, C., Bell, S. & Bradley, A. Influence of urban edges on stress in an arboreal mammal: a case study of squirrel gliders in southeast Queensland, Australia. *Landscape ecology* **27**, 1407–1419, <https://doi.org/10.1007/s10980-012-9790-8> (2012).
63. Roth, L. S., Faresjo, A., Theodorsson, E. & Jensen, P. Hair cortisol varies with season and lifestyle and relates to human interactions in German shepherd dogs. *Sci Rep* **6**, 19631, <https://doi.org/10.1038/srep19631> (2016).
64. Henley, P. *et al.* Cultural and socio-economic conditions as factors contributing to chronic stress in sub-Saharan African communities. *Canadian Journal of Physiology and Pharmacology* **92**, 725–732, <https://doi.org/10.1139/cjpp-2014-0035> (2014).
65. Gidlow, C. J., Randall, J., Gillman, J., Smith, G. R. & Jones, M. V. Natural environments and chronic stress measured by hair cortisol. *Landscape and Urban Planning* **148**, 61–67 (2016).
66. Natt, D., Johansson, I., Faresjo, T., Ludvigsson, J. & Thorsell, A. High cortisol in 5-year-old children causes loss of DNA methylation in SINE retrotransposons: a possible role for ZNF263 in stress-related diseases. *Clinical Epigenetics* **7**, <https://doi.org/10.1186/s13148-015-0123-z> (2015).
67. Gonzalez-de-la-Vara, M. D. *et al.* Effects of adrenocorticotrophic hormone challenge and age on hair cortisol concentrations in dairy cattle. *Canadian Journal of Veterinary Research-Revue Canadienne De Recherche Veterinaire* **75**, 216–221 (2011).
68. Jarcho, M. R., Massner, K. J., Eggert, A. R. & Wichelt, E. L. Behavioral and physiological response to onset and termination of social instability in female mice. *Horm Behav* **78**, 135–140, <https://doi.org/10.1016/j.yhbeh.2015.11.004> (2016).
69. Bryan, H. M., Adams, A. G., Invik, R. M., Wynne-Edwards, K. E. & Smits, J. E. Hair as a meaningful measure of baseline cortisol levels over time in dogs. *Journal of the American Association for Laboratory Animal Science* **52**, 189–196 (2013).
70. Chan, J., Sauvé, B., Tokmakejian, S., Koren, G. & Van Uum, S. Measurement of cortisol and testosterone in hair of obese and non-obese human subjects. *Experimental and clinical endocrinology & diabetes: official journal, German Society of Endocrinology [and] German Diabetes Association* **122**, 356–362 (2014).
71. Lightman, S. L. & Conway-Campbell, B. L. The crucial role of pulsatile activity of the HPA axis for continuous dynamic equilibration. *Nat. Rev. Neurosci.* **11**, 710 (2010).
72. Whitten, P., Brockman, D. & Stavisky, R. Recent advances in noninvasive techniques to monitor hormone-behavior interactions. *Am J Phys Anthropol* **107**, 1–23 (1998).
73. Boesch, M. *et al.* Hair cortisol concentration is unaffected by basic military training, but related to sociodemographic and environmental factors. *Stress* **18**, 35–41 (2015).
74. Gidlow, C. J., Randall, J., Gillman, J., Silk, S. & Jones, M. V. Hair cortisol and self-reported stress in healthy, working adults. *Psychoneuroendocrinology* **63**, 163–169 (2016).
75. Campbell, J. & Ehlert, U. Acute psychosocial stress: does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology* **37**, 1111–1134 (2012).
76. Gao, W. *et al.* Temporal features of elevated hair cortisol among earthquake survivors. *Psychophysiology* **51**, 319–326, <https://doi.org/10.1111/psyp.12179> (2014).
77. Heinze, K., Lin, A., Reniers, R. L. & Wood, S. J. Longer-term increased cortisol levels in young people with mental health problems. *Psychiatry Res* **236**, 98–104, <https://doi.org/10.1016/j.psychres.2015.12.025> (2016).

78. Steudte-Schmiedgen, S. *et al.* Hair cortisol concentrations and cortisol stress reactivity predict PTSD symptom increase after trauma exposure during military deployment. *Psychoneuroendocrinology* **59**, 123–133, <https://doi.org/10.1016/j.psyneuen.2015.05.007> (2015).
79. Chu, X.-X. *et al.* A natural model of behavioral depression in postpartum adult female cynomolgus monkeys (*Macaca fascicularis*). *Zoological Research* **35**, 174–181 (2014).
80. Klumbies, E., Braeuer, D., Hoyer, J. & Kirschbaum, C. The Reaction to Social Stress in Social Phobia: Discordance between Physiological and Subjective Parameters. *Plos One* **9**, <https://doi.org/10.1371/journal.pone.0105670> (2014).
81. Kapoor, A., Lubach, G. R., Ziegler, T. E. & Coe, C. L. Hormone levels in neonatal hair reflect prior maternal stress exposure during pregnancy. *Psychoneuroendocrinology* (2016).
82. Skoluda, N., Dettenborn, L., Stalder, T. & Kirschbaum, C. Elevated hair cortisol concentrations in endurance athletes. *Psychoneuroendocrinology* **37**, 611–617 (2012).
83. Dettenborn, L., Tietze, A., Bruckner, F. & Kirschbaum, C. Higher cortisol content in hair among long-term unemployed individuals compared to controls. *Psychoneuroendocrinology* **35**, 1404–1409, <https://doi.org/10.1016/j.psyneuen.2010.04.006> (2010).
84. Hunt, K. E. *et al.* Baleen hormones: a novel tool for retrospective assessment of stress and reproduction in bowhead whales (*Balaena mysticetus*). *Conservation physiology* **2** (2014).
85. Hunt, K. E., Lysiak, N. S., Moore, M. & Rolland, R. M. Multi-year longitudinal profiles of cortisol and corticosterone recovered from baleen of North Atlantic right whales (*Eubalaena glacialis*). *Gen. Comp. Endocrinol.* **254**, 50–59 (2017).
86. Kapoor, A., Schultz-Darken, N. & Ziegler, T. E. Radiolabel validation of cortisol in the hair of rhesus monkeys. *Psychoneuroendocrinology* **97**, 190–195 (2018).
87. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & Group, P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* **6**, e1000097 (2009).

Acknowledgements

Academic librarian Kjersti Aksnes-Hopland, University of Bergen Library, assisted in the development of search strategy and search implementation. We would furthermore like to extend our gratitude to the authors who took the time to respond to our queries and supplied missing information.

Author Contributions

The authors – Otto Kalliokoski, Finn Jellestad and Robert Murison – contributed equally to design, execution, analysis, and reporting of the study.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-48517-2>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019