RESEARCH

Validation of a commercial enzyme immunoassay to assess urinary oxytocin in humans

Franka S Schaebs[1,2], Gwen Wirobski[3], Sarah Marshall-Pescini [ID][3]**, Friederike Range[3]** and **Tobias Deschner[1]**

[1]Interim Group Primatology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Saxony, Germany
[2]ZLS, University of Leipzig, Leipzig, Saxony, Germany
[3]Domestication Lab, Wolf Science Center, Konrad Lorenz Institute of Ethology, University of Veterinary Medicine Vienna, Vienna, Austria

Correspondence should be addressed to G Wirobski or T Deschner: Gwendolyn.Wirobski@vetmeduni.ac.at or deschner@eva.mpg.de

Abstract

Within the last decade, oxytocin (OT) has attracted a lot of attention in the context of various human social behaviors. Besides its importance in regulating physiological processes in females related to giving birth and lactation, OT is involved in the establishment and maintenance of social relationships, trust and emotion recognition. However, results are not always consistent across studies, which may partly be due to the incomplete validation of methods used to assess OT levels. Carefully validating a method before its use is of crucial importance to ensure that it can be used to accurately, reliably and repeatedly assess OT levels. With this study we evaluated a commercially available Enzyme Immunoassay to assess OT in human urine samples by conducting a careful analytical validation. Results indicate that, with regard to parallelism and immunoreactivity, human urinary OT can be assessed reliably. However, extraction methods need further improvement to optimize measures of accuracy and extraction efficiency, especially in the lower range of the assay system. Tests on OT stability indicate that OT is affected by degradation when stored at 4°C or room temperature. Storing urine samples over longer periods revealed that OT levels are most stable when stored as ethanol extracts at −20°C compared to being stored as samples at −20°C or −80°C. Although some of the validated parameters did not reach the intended quality criteria, this study highlights the importance of such in depth validation procedures and reporting results to make them available to researchers embarking on projects utilizing such methods.

Introduction

Recently, numerous studies have investigated the role of the oxytocinergic system in human social behavior with results showing that it plays a role in bonding (1), parent-infant relationship (2), trust (3, 4), emotion recognition (4, 5) and stress (6, 7), suggesting that the oxytocinergic system played an important role in the evolution of human behavior (8). However, results are often inconsistent across studies, questioning the robustness of the effects. In addition, a number of authors have suggested that a 'publication bias' may have led to an over-representation of positive results in the literature (9, 10, 11).

Oxytocin (OT), a nine amino-acid neuropeptide hormone, is produced centrally in the hypothalamus and released into the blood stream via the pituitary. Many studies have assessed the function of central OT in mammalian species (e.g. primates: (12, 13, 14); rodents: (15, 16, 17, 18)), including humans (19, 20, 21). Since central measures of OT are highly invasive, peripheral

measures of OT in blood, urine and saliva are more often used in studies investigating its role in mediating human behavior. However, aside from the lively debate regarding what extent peripheral OT measures reflect central mechanisms which are more likely involved in mediating behavior (22, 23, 24, 25) there has also been a lack of in-depth validation procedures of peripheral measures prior to their use in behavioral studies.

Numerous published studies, rather than providing a method validation, often refer to other studies supposedly providing validation of their extraction method and assay system, which however, on closer inspection, do not report the details of their findings (reviewed in (24)). In other cases, authors refer to published validation studies, but then implement changes or even omit specific procedures in their own methods (reviewed in (24, 25)), which may lead to non-replicate findings or hinder an appropriate interpretation of results.

As excretion routes and patterns may differ tremendously depending on the matrix at hand, it is of crucial importance to conduct a solid analytical validation for each single type of sample (e.g. urine or blood). Equally, different extraction methods and assay systems must be validated in order to confirm that in fact one measures the biological marker itself or its metabolites and not some other immunoreactive substance present in the samples, which then would lead to incorrect findings and interpretations.

Several steps must be included in a complete analytical validation. First of all, a test for parallelism should be conducted to assess the potential presence of matrix effects, which could interfere with the assay system leading to a pattern where, depending on the concentration of the hormone in the samples, high levels might be underestimated, whereas low levels might be overestimated or vice versa. Second, to evaluate the extraction method and assay system used, measures of their validity should be included. To assess patterns of immunoreactivity of the respective hormone or its metabolites in the specimen and species of interest, an immunogram should be conducted. Finally, to account for variance in levels of the hormone of interest depending on how and for how long samples are stored before further processing in the lab, reasonable storage conditions should be tested for their validity.

In the present study, we aimed at validating a method to non-invasively assess measures of peripheral OT in human urine samples, using a commercially available Enzyme Immunoassay (Enzo Life Sciences, Assay designs, Cat. No. 901-153A-0001). To do so,

we conducted a test for parallelism to assess the potential presence of matrix effects. We provide measures to assess the efficacy of the extraction method and assay system as well as an immunogram to investigate patterns of immunoreactivity in human urine samples. Furthermore, to be able to evaluate the stability of OT in human urine samples we conducted a test of different long- and short-term storage conditions. This in depth validation procedure aimed to provide a reliable tool to assess levels of urinary OT in humans for our own and other researchers' future studies.

Methods

Subjects

Urine samples of nine subjects (six women, three men) were collected at the Max-Planck-Institute for Evolutionary Anthropology (MPI EVA) in Leipzig, Germany. Urine samples of women were only collected when no menstrual bleeding was recorded. Ethical approval was obtained from the ethical commission of the Max Planck Society (Approval number: 2017_09). Consent has been obtained from each subject after full explanation of the purpose and nature of all procedures used.

Sample collection

Samples were collected with plastic trays (Carl Roth, CEN 7.1). Urine samples were acidified by adding 100 µL of a 0.1% phosphoric acid to each 1 mL aliquot (26, 27) to avoid degradation of oxytocin and all samples were frozen directly at −20°C.

Sample preparation and urinary oxytocin measurements

Laboratory analyses were conducted in the Endocrinology laboratory at the MPI EVA and urine samples were stored at −20°C until used for validation. All urine samples were extracted using a solid phase extraction (SPE) following a previously published protocol (26), with some minor modifications (28). To assess immunoreactive urinary oxytocin metabolites (iuOTM), extracts were measured in duplicates using a commercially available assay system by Enzo Life Sciences designed to measure oxytocin levels (Cat. No. 901-153A-0001). In case OD values differed more than 10%, measures were repeated or excluded from further analysis.

Inter-assay coefficients of variance were 9.5% for a high concentrated OT standard and 17.6% for a low concentrated OT standard ($n = 15$). In the current study, the low concentrated OT standard was close to the lower end of the linear range of the assay, which has produced a high CV. Thus, samples falling within the lower range of the assay may be associated with considerable variation calling for careful interpretation as subtle differences in magnitude (effect sizes) may not be detected reliably. Intra-assay coefficient of variance, calculated via the average variability across duplicates of 31 samples measured in a single assay, was 6.9% ($n = 34$).

Analytical validation

Parallelism
Matrix effects, that could potentially interfere with the assay system, were investigated by conducting a test for parallelism. A 450 µL pool sample was spiked with 50 µL of a 4000 pg/mL OT standard (supplied by Enzo Life Sciences, Assay designs) and diluted serially. As the concentration of iuOTM in human urine samples can be low, pool samples were spiked with an OT standard. This was done to obtain a meaningful dilution curve that allowed to assess matrix effects at concentrations, samples are routinely measured at (28, 29).

Extraction efficiency and assay accuracy
To assess extraction efficiency, we created five pools of urine samples and spiked 237.5 µL of each pooled sample with 12.5 µL of three different concentrations of an OT standard (high = 4000 pg/mL; medium = 2000 pg/mL; low = 1000 pg/mL) preceding extraction. Furthermore, to assess assay accuracy, 237.5 µL of each pooled sample were spiked with 12.5 µL of three different concentrations of an OT standard (high = 4000 pg/mL; medium = 2000 pg/mL; low = 1000 pg/mL) following extraction. Values from the spiked and unspiked samples were used to calculate percent recovery for extraction efficiency and assay accuracy (28, 30).

Immunograms
Patterns of immunoreactivity were investigated by creating an immunogram of human urine samples. For this, we created a pooled urine sample and a 2000 pg/mL

OT standard. Both were extracted as described before and each reconstituted in 150 µL 30% ACN. Twelve fractions were separated (using an Alliance 2695 HPLC, Gemini C18 column, Phenomenex) and collected (using a Waters Fraction Collector 3, Waters) for each, the standard and the pooled urine sample (see (28) for details). Following lyophilization, samples were kept at −20°C until further analysis for immunoreactivity (IR) in each of the fractions using the Enzo immunoassay. Proportions of explained and unexplained IR in each of the fractions were calculated as given in a previous publication (28).

Storage and stability
In order to evaluate degradation patterns of iuOTM levels, we conducted a set of three different tests. First, the effect of short-term storage on urine samples at either 4°C or room temperature (RT) was assessed by creating two pooled urine samples. Aliquots of the pooled urine samples were stored directly and 1, 3, 6 and 24 h following collection.

Second, the effect of repeated freezing and thawing was assessed by exposing two pooled urine samples to three repeated freezing and thawing cycles. So far, studies added phosphoric acid to urine samples to prevent OT degradation (26, 27). Therefore, to evaluate the effect of phosphoric acid, we tested each condition of the first two sets (short-term storage and repeated freezing and thawing) with phosphoric acid (PA) added to the aliquots or without adding acid.

Third, the effect of long-term storage was assessed by storing four urine samples either directly at −20°C or at −80°C or as extracts stored in 300 µL 100% EtOH for 2 weeks, 4 weeks, 12 weeks and 6 months.

Statistical analysis

All plots were created using R (version 3.3.3; (31)). We tested for parallelism by fitting a linear model including the interaction between sample type (standard curve and pooled sample) and the concentration of the standard with the percent binding as response variable (28, 29). The model was fitted in R (31) using the function lm. The check for assumptions of normality and homogeneity of the residuals did not indicate any problems (inspection of a qq-plot of the residuals and residuals plotted against fitted values (32). Model stability was assessed by means of DFBeta (32), which did not indicate any problems.
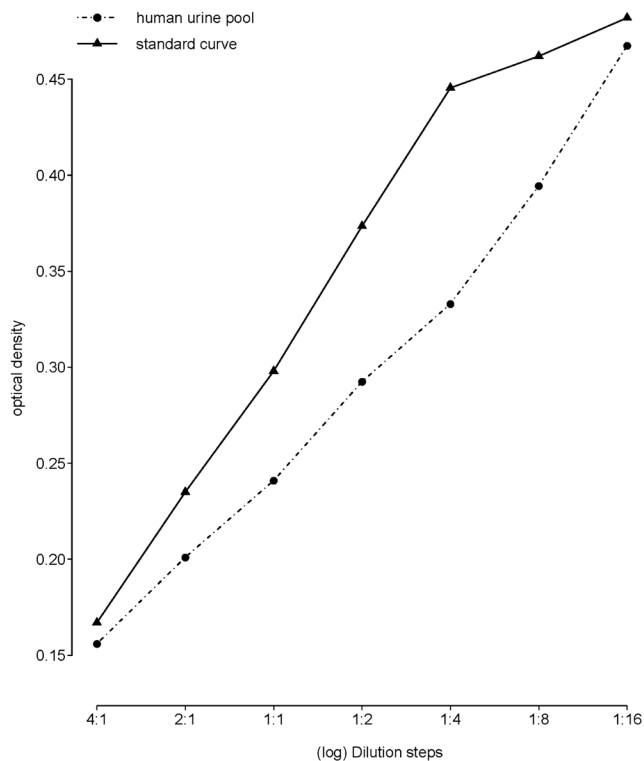
Figure 1
Parallelism of a serially diluted human urine sample pool to the standard curve ($t_{(10)} = -0.507$, $P = 0.623$). Note that the x-axis is on a log scale.

Results

Parallelism

The serially diluted human pool sample was parallel to the standard curve ($t_{(10)} = -0.507$, $P = 0.623$), which was confirmed by visual inspection (Fig. 1).

Extraction efficiency and assay accuracy

Mean extraction efficiency was calculated for spiked human urine pool samples. When spiked with a high concentrated OT standard the mean extraction efficiency was 92.8% (range: 71.9–120%, SD = 19.4, $n = 5$). A medium spiking resulted in 98.9% (range: 66.9–131%, s.d. = 23.2, $n = 5$) and low spiking revealed a 101% (range: −16.2 to 156, s.d. = 67.82, $n = 5$) extraction efficiency.

Mean assay accuracy was also calculated for spiked human urine pool samples. When spiked with a high concentrated OT, standard mean assay accuracy was 126% (range: 108–160%, s.d. = 21.4, $n = 5$). With a medium concentrated OT, it was 120% (range: 104–140%, s.d. = 14.8, $n = 5$) and with a low concentrated OT, it was 113% (range: 74.8–190%, s.d. = 47.0, $n = 5$).
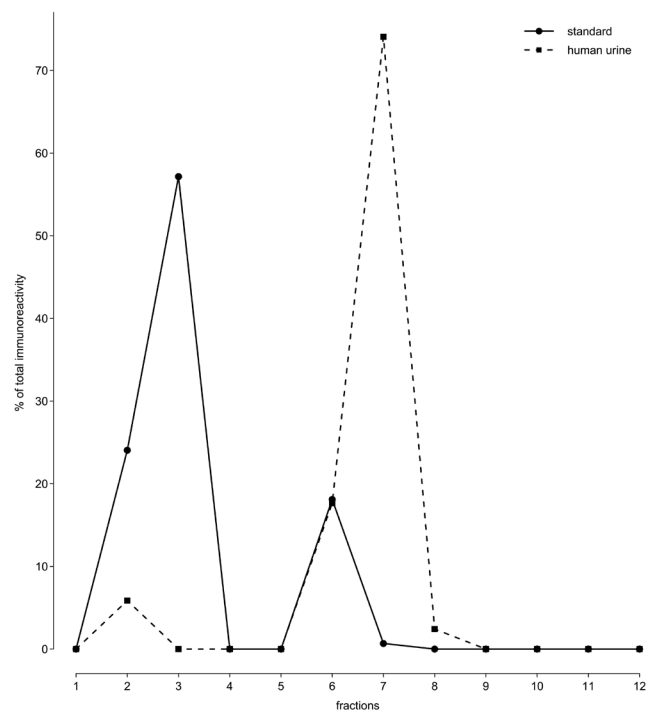

Figure 2
Immunoreactivity in extracted human urine and OT standard.

Immunograms

The immunogram of an extracted OT standard showed immunoreactivity in fractions 2, 3, 6 and 7. These accounted respectively for 24.1, 57.2, 18.1 and 0.7% of the total immunoreactivity (Fig. 2). The immunogram of the extracted human urine showed immunoreactivity in fractions 2, 6, 7 and 8. These accounted, respectively, for 5.9, 17.7, 74.1 and 2.4% of the total immunoreactivity (Fig. 2). Therefore, 97.6% of immunoreactivity in extracted human urine is accounted for by the immunoreactivity found in extracted standard.

Storage and stability

Short-term storage before freezing samples

When PA was added to the urine samples, we observed an initial average increase in iuOTM of 105% (range: −23.6% to +234%) 1 h following collection and storage at 4°C. After that, levels of iuOTM decreased again and stayed constant for a storage of up to 24 h with an average increase of 41.3% (range: −41.9% to +126%) as compared to the freshly collected urine samples (Fig. 3). If no PA was added to the urine samples following collection, iuOTM levels initially increased on average 24.8% within the first 3 h (range: −34.5% to +93.9%) compared to the freshly collected urine
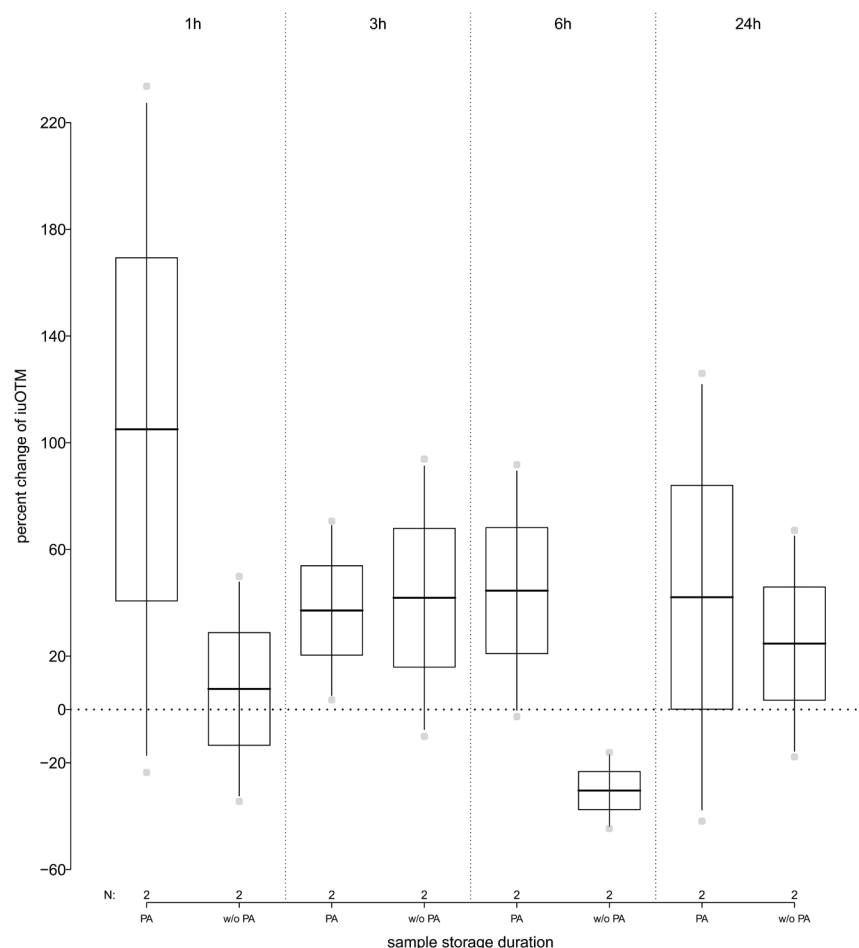
Figure 3
Percent change in immunoreactive urinary oxytocin metabolite (iuOTM) levels of human urine pool samples (*n* = 2) stored at 4°C for 1, 3, 6 and 24 h; separated by samples to which phosphoric acid (PA) was added or not (w/o PA). The horizontal line indicates the 100% control iuOTM levels, for pooled samples that have been frozen immediately following collection.

samples. After 6 h of storage at 4°C, iuOTM levels were on average 30.4% (range: −16.1% to 44.7%) less as compared to the original measurement, but then increased again to on average 24.7% (range: −17.8% to 67.1%) more when stored for up to 24 h, as compared to the freshly collected urine samples (Fig. 3 and see Table 1 for details). Storing samples at room temperature for up to 24 h revealed iuOTM levels to be less stable in comparison to storage at 4°C.

When PA was added to freshly collected urine samples stored at RT, iuOTM levels initially increased on average 57.3% (range: −27.2% to +156%) when stored for up to 3 h. They decreased by on average 17.0% (range: −32.5% to +66.5%) when stored for up to 6 h and increased again to on average 68.8% (range: −27.1% to 165%) when stored for up to 24 h as compared to the freshly collected urine samples. When no PA was added following collection and samples were stored at RT, iuOTM levels increased on average by 35.6% (range: −25.7% to +96.8%), when stored for 1 h and then decreased to an average of −24.3% (range −46.3% to 2.3%) as compared to the freshly collected

Table 1 Short-term storage of human urine samples before freezing of samples at 4°C or at room temperature, with either phosphoric acid added following collection or no PA added. Data shown is average (and range) percent change of the iuOTM levels as compared to the original measurement directly following collection.

Storage method	Storage duration			
	1 h	3 h	6 h	24 h
4°C, PA added	106% (−23.6% to +234%)	37.1% (+3.6% to +70.6%)	44.6% (−2.7% to +91.7%)	42.1% (−41.9% to +126%)
4°C, No PA added	7.7% (−34.5% to +49.9%)	41.9% (−10.1% to +93.9%)	−30.4% (−44.7% to −16.4%)	24.7% (−17.8% to +67.1%)
RT, PA added	64.4% (−27.2% to +156 %)	50.3% (−17.5% to +118%)	17.0% (−32.5% to +66.5%)	68.8% (−27.2% to +165%)
RT, No PA added	35.6% (−25.7% to +96.9%)	13.7% (+0.1% to +27.3%)	−24.3% (−46.3% to −2.3%)	21.1% (−26.9% to +68.9%)

PA, phosphoric acid; RT, room temperature.
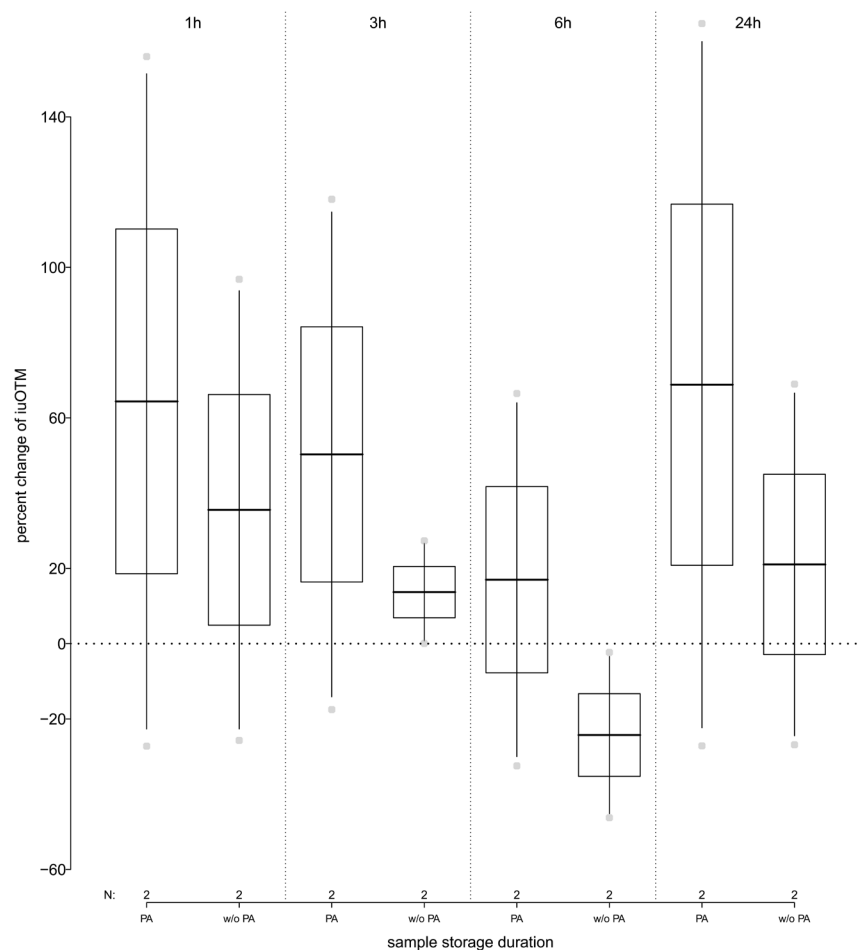
Figure 4
Percent change in immunoreactive urinary oxytocin metabolite (iuOTM) levels of human urine pool samples (*n* = 2) stored at room temperature (RT) for 1, 3, 6 and 24 h; separated by samples to which phosphoric acid (PA) was added or not (w/o PA). The horizontal line indicates the 100% control iuOTM levels, for pooled samples that have been frozen immediately following collection.

urine samples, when stored for up to 6 h. IuOTM levels increased again to on average plus 21.1% (range: −21.9% to +68.9%) when stored for up to 24 h (Fig. 4 and Table 1 for details).

Exposure of human pooled urine samples to repeated thawing

When PA was added immediately to the samples following collection, iuOTM levels stayed constant when thawed up to two times with an average decrease of 2.7% (range: −14.9% to +6.9%) and increased above control levels (average: +35.7%; range: +14.4% to +56.9%). When no PA was added, iuOTM levels stayed relatively stable when samples were thawed up to three times (average: +1.4%; range: −14.2% to +14.4%) (Fig. 5 and see Table 2 for details).

Long term storage

Overall, iuOTM levels of human urine samples increased with increasing duration of storage, independent of the

storage condition. After a storage duration of 6 months, iuOTM levels increased on average by 150% (range: +68.6% to 250%) when stored as samples at −20°C, by 79.2% (range: +33.6% to 138%) when stored as samples at −80°C and by 52.9% (range: +17.5% to +77.5%) when extracted preceding storage and stored as ethanol extracts at −20°C (Fig. 6).

Discussion

Here we provide a detailed analytical validation to assess iuOTM levels in human urine samples using a commercially available Enzyme Immunoassay provided by Enzo Life Sciences. Even though the company developing this assay states that it can be used for measurements of OT in human urine, serum, plasma, breast milk, saliva, and cerebrospinal fluid samples, our study indicates that there are considerable issues when measuring OT in urine, especially with its accuracy. This should be considered, when interpreting the results of previously published work that has used this assay (reviewed in (33)).
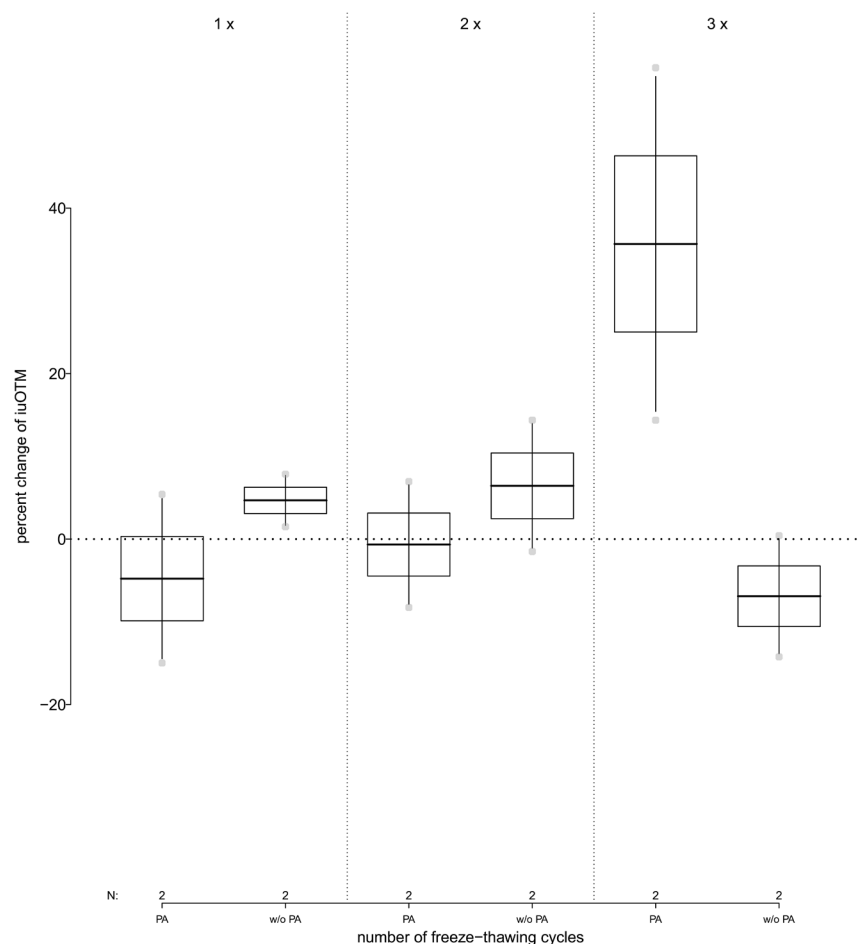
Figure 5
Percent change of immunoreactive urinary oxytocin metabolite (iuOTM) levels in human urine pool samples ($n = 2$) over the course of three freeze-thawing cycles; separated for samples to which phosphoric acid (PA) was added or not (w/o PA). The horizontal line indicates the 100% control iuOTM level, which was not exposed to thawing.

When tested for parallelism, results indicate that matrix effects do not appear to be a problem and iuOTM can be reliably measured in human urine samples. However, with regard to extraction efficiency and assay accuracy, we offer a cautionary note. Both measures revealed standard deviations that are rather high, but still acceptable when spiked with a high or medium concentrated standard. However, when samples are spiked with a low concentrated standard, standard deviations were unacceptably high. This indicates that care must be taken, especially when interpreting OT measures in the lower range of the assay. Thus, to determine the suitability of sample substrate and

mode of measurement (in this case: urine, EIA) for one's study purpose, it is crucial to consider two points: (i) The expected proportion of samples which will fall below a certain concentration threshold and thus may be associated with large variability/uncertainty, and (ii) the expected effect size to be measured (i.e. the expected magnitude of difference between samples). If only very few samples are expected to fall within the lower range of the assay, they will most likely not affect overall patterns of results and, thus, concerns about their limited accuracy may not be so grave. Similarly, if the expected magnitude of difference between samples is expected to be larger than the variability

Table 2 Exposure of human urine samples to repeated thawing, with either phosphoric acid added following collection or no phosphoric acid added. Data shown is average (and range) percent change of the iuOTM levels as compared to the original measurement directly following collection.

	Thawing cycles		
Storage method	1×	2×	3×
PA added	−4.8% (−15.0% to +5.4%)	−0.7% (−8.3% to +6.9%)	35.7% (+14.4% to +56.9%)
No PA added	4.1% (+1.5% to +7.8%)	6.4% (−1.5% to +14.4%)	−6.9% (−14.2% to +0.4%)

PA, phosphoric acid.
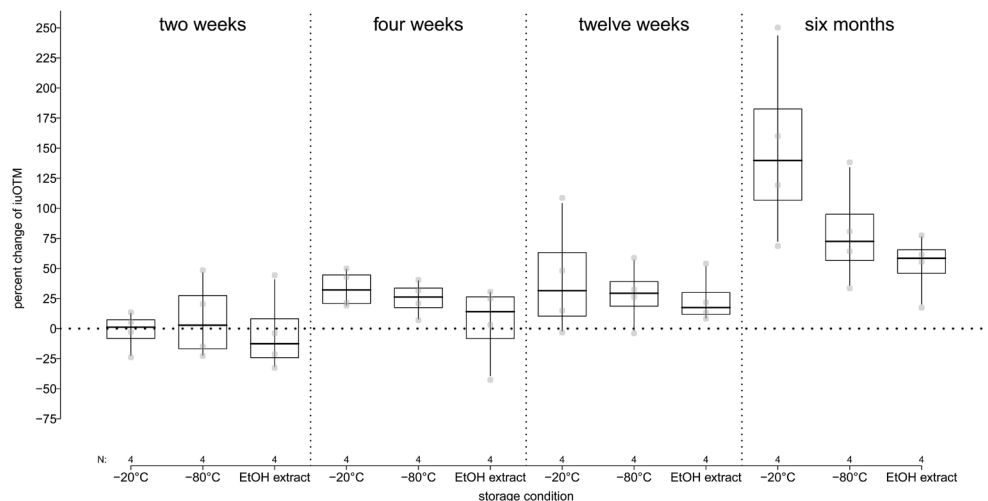
Figure 6

Percent change in immunoreactive urinary oxytocin metabolite (iuOTM) levels of four human urine samples after storage as sample at either −20°C or −80°C or as ethanol (EtOH) extract for 2 weeks (corrected for QC), 4 weeks, 12 weeks and 6 months. The horizontal line indicates the 100% control iuOTM levels (i.e. zero change).

associated with the measures, then valid conclusions may still be drawn from the data if such differences were to be found. In addition, future studies should address the need for further improvement of urine extraction protocols, and, above all, transparently report assay parameters to allow readers to reach their own conclusions regarding the potential issues with specific measurements.

Comparing results to published studies, it appears that measures of OT are less reliable with regard to recovery measures than those of other biological markers ([30], [34], [35], [36]), but comparable to what was reported for OT in dog and wolf urine using the same extraction procedure and assay system ([28]), including similar limitations with regard to accuracy in the low-concentration range as reported in an in-depth validation of plasma OT measures of seven domestic species using the same extraction procedure and assay system ([37]). However, a few studies, including one on human urine samples, have partially validated the measurement of OT by conducting tests for parallelism and accuracy (as summarized in 33, using an EIA by Assay Designs) and reported acceptable results. Though, results on extraction efficiency and details on how the samples were processed (spiked, etc.) are missing.

Future studies should first investigate improved extraction methods by, for example, testing various solid phase cartridges and extraction solutions, thereby optimizing results of iuOTM levels. Second, more effort should be made in reporting not only results, but also detailed method descriptions and validation parameters. This is important as only if such details are given, the (naïve) reader as well as reviewers are able to evaluate the

quality, accuracy and the relevance of the respective study. In addition, only if such details are given is it possible to evaluate whether results are comparable across studies.

Immunograms

To identify if the assay system that is used does in fact measure the hormone of interest, one should assess patterns of immunoreactivity in the samples at hand. In the present study, we found that immunoreactivity in human urine samples was present in several fractions with two peaks. However, the majority of the fractions in which immunoreactivity was detected in the pool sample were identical with the fractions in which immunoreactivity was detected in the standard. Therefore almost 100% of the immunoreactivity can be explained by immunoreactivity found in an extracted OT standard.

Ideally, in fractions of a standard, immunoreactivity should be found in one fraction. As OT is reportedly susceptible to degradation caused by various factors such as increased temperatures and pH-level ([38]), steps undertaken during the extraction and handling process might lead to cleavage of functional groups of the OT molecule. If the assay picks up immunoreactivity in several fractions of the standard, these metabolites are related to OT and can be considered immunoreactive OT metabolites (iOTM).

Our results show that although almost all of the immunoreactivity found in human urine samples is congruent with that found in the OT standard, one peak did not coincide with the standard. Such a pattern has

already been found for extracted human urine samples (23) and extracted plasma samples of prairie voles (39). However, another study reported one single peak of immunoreactivity in extracted human urine that was matching the immunoreactivity in a synthetic OT standard (40). None of these studies reported whether the OT standard was extracted or not, which might be a factor that perhaps could contribute to observed differences. In addition, a previous study of our group found the same pattern when OT immunoreactivity was assessed in urine samples of dogs and wolves (28). In that study we also collected fractions of an unextracted OT standard and found that immunoreactivity was present in several fractions with one peak (as compared to two peaks in extracted OT standard), which supports the hypothesis of structural changes of the OT molecule during extraction and handling. Further investigation of how OT is metabolized within subjects, using radiometabolism studies, would help to identify OT metabolites. Using a radiometabolism study allows to differentiate between immunoreactivity that is caused by OT and its metabolites or substances unrelated to OT. By injecting a radioactively labeled hormone in an individual and repeatedly collecting the samples of interest, for example urine samples, one can track routes and patterns of excretion. If then samples are measured with an HPLC, immunoreactivity can be assessed in collected fractions and thus specific OT metabolites or substances not related to OT can be identified. In a previous study (41), fractions of radioactively labeled urine samples of common marmosets were collected and measured and it was found that only parts of the radioactivity found in urine samples was congruent with an unextracted OT standard. The authors suggest this to be due to a metabolic breakdown of the OT molecule or deposit in other body tissues (which has been suggested before, see (42)). However, such results cannot easily be generalized as metabolism differs across species and this would have to be conducted for each species separately. Furthermore, conducting a radiometabolism study requires time consuming procedures and is less feasible in humans due to its high invasiveness.

Storage and stability

As hormones, especially in urine and fecal samples, are prone to bacteria, they are likely exposed to degradation. This can have an impact on levels of hormones and their metabolites. Thus, when assessing hormonal levels in peripheral specimen, it becomes very important to identify degradation patterns during handling and storage and

to evaluate various storage methods for samples before analysis.

Here we exposed human urine samples to various short- (storage at 4°C and room temperature, as well as repeated thawing) and long-term (storage of samples at −20°C or −80°C, or as extracts at −20°C) storage conditions.

Short-term storage of human urine samples at 4°C or room temperature

While in general steroid hormones are thought to be stable when stored at room temperature for about 24 h and for some days at 4°C, freezing of samples as soon as possible following collection is generally recommended (43, 44). Other biological markers, such as peptide hormones or gonadotropins, are more susceptible to degradation and besides freezing samples as soon as possible, adding a preservative to prevent degradation is recommended (33, 43).

In the present study, levels of iuOTM in human urine samples being stored for a short-term period at either 4°C or at room temperature (RT) revealed quite a degree of variability. However, iuOTM levels appear to be more stable when phosphoric acid (PA) is added immediately following collection of the samples. Thus, to prevent degradation of iuOTM levels in human urine samples, we suggest to add PA to freshly collected samples and prepare them for long-term storage if needed or process samples for further analysis as soon as possible.

Exposure of human pooled urine samples to repeated thawing

Even though many biological markers are generally stable over repeated thawing, this should be avoided if possible (33, 45, 46). However, repeated thawing of samples can easily occur when samples are transported from the field to the lab or during repeated handling in the lab. If that is the case, degradation may occur due to the presence of bacteria.

So far results on the effect of repeated thawing on OT vary. While one study reported lower levels of urinary OT in humans after three thawing cycles (47), another study did not find such an effect on human blood samples after five thawing cycles (23). Here we found that iuOTM levels were stable when thawed twice (for both, PA and non-PA samples), remained stable in samples where no PA was added for up to three times thawing, but increased when thawed three times in samples to which PA was added.

These results indicate that repeated thawing has a minor effect on iuOTM levels, but periods at room temperature or 4°C should be kept as short as possible and if applicable should be reported in publications. In general, samples should be aliquoted immediately following collection to avoid repeated thawing due to handling procedures.

Long-term storage of human urine

Often samples have to be stored following collection before being analyzed for the hormone of interest, especially when samples are collected under field conditions. As previous work has already shown variation in OT levels (40) in relation to storage duration, it becomes important to assess degradation patterns over longer periods of time and to identify the best possible storage condition. We exposed samples to three different conditions and stored them as samples at −20°C and −80°C and as ethanol extracts at −20°C. Results show that overall, iuOTM levels increase with increasing storage duration and that iuOTM levels are most stable when stored as ethanol extracts. Given our results, we recommend storage of samples as ethanol extracts until being processed for further analysis. If this is not feasible, samples should be stored at −80°C and extracted as soon as possible. Batches of samples stored for the same duration, should be measured for the hormone of interest in one run. In case samples have to be collected under field conditions, they should be stored in liquid nitrogen until transport to a lab and further processing (27, 48, 49). In all cases, storage condition and duration should be reported in publications to allow the reader to estimate its potential effect.

Conclusion and general considerations

With this study we have shown that OT in human urine samples can be assessed using a commercially available Enzyme Immunoassay produced by Enzo Life Sciences. However, some of the quality parameters achieved are not optimal. Precisely for that reason it is of major importance that parameters for such analytical validations are reported in publications in order to evaluate their significance and relevance both for already published studies and help inform choices for future studies. Even if a validation shows that results are far from optimal, results should be reported (along with the results of the respective study) so that reviewers and readers can retrace each methodological step and then evaluate their importance in affecting the overall pattern of results.

Given the results of the present study, the assay system used here (Enzo Life Sciences) can be used to assess OT in human urine samples, but there is need to improve extraction methods. To do so, different solid phase cartridges and extraction solutions could be tested for their accuracy and efficiency and results should be compared. Furthermore, the possibility of choosing alternative assay systems, for example the OT Arbor assay, should be taken into consideration.

In addition, care must be taken when choosing the storage method as OT is prone to degradation. Samples should be extracted as soon as possible following collection (and stored at −80°C until extraction) and extracts stored at −20°C until measured. Ideally, studies should be planned ahead so that all samples are stored for about the same period of time and under the same conditions, thereby improving comparability across samples.

References

1 Gordon I, Zagoory-Sharon O, Schneiderman I, Leckman JF, Weller A & Feldman R. Oxytocin and cortisol in romantically unattached young adults: associations with bonding and psychological distress. *Psychophysiology* 2008 **45** 349–352. (https://doi.org/10.1111/j.1469-8986.2008.00649.x)

2 Feldman R, Gordon I & Zagoory-Sharon O. Maternal and paternal plasma, salivary, and urinary oxytocin and parent-infant synchrony: considering stress and affiliation components of human bonding. *Developmental Science* 2011 **14** 752–761. (https://doi.org/10.1111/j.1467-7687.2010.01021.x)

3 Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U & Fehr E. Oxytocin increases trust in humans. *Nature* 2005 **435** 673–676. (https://doi.org/10.1038/nature03701)

4 Van IJzendoorn MH & Bakermans-Kranenburg MJ. A sniff of trust: meta-analysis of the effects of intranasal oxytocin administration on face recognition, trust to in-group, and trust to out-group. *Psychoneuroendocrinology* 2012 **37** 438–443. (https://doi.org/10.1016/j.psyneuen.2011.07.008)

5 Shahrestani S, Kemp AH & Guastella AJ. The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: a meta-analysis. *Neuropsychopharmacology* 2013 **38** 1929–1936. (https://doi.org/10.1038/npp.2013.86)

6 Eckstein M, Scheele D, Weber K, Stoffel-Wagner B, Maier W & Hurlemann R. Oxytocin facilitates the sensation of social stress. *Human Brain Mapping* 2014 **35** 4741–4750. (https://doi.org/10.1002/hbm.22508)

7 Valstad M, Alvares GA, Egknud M, Matziorinis AM, Andreassen OA, Westlye LT & Quintana DS. The correlation between central and peripheral oxytocin concentrations: a systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews* 2017 **78** 117–124. (https://doi.org/10.1016/j.neubiorev.2017.04.017)

8 Carter CS. Oxytocin pathways and the evolution of human behavior. *Annual Review of Psychology* 2014 **65** 17–39. (https://doi.org/10.1146/annurev-psych-010213-115110)

9 Walum H, Waldman ID & Young LJ. Statistical and methodological considerations for the interpretation of intranasal oxytocin studies. *Biological Psychiatry* 2016 **79** 251–257. (https://doi.org/10.1016/j.biopsych.2015.06.016)

10 Lane A, Luminet O, Nave G & Mikolajczak M. Is there a publication bias in behavioural intranasal oxytocin research on humans? Opening the file drawer of one laboratory. *Journal of Neuroendocrinology* 2016 **28** [epub]. (https://doi.org/10.1111/jne.12384)

11 Leng G & Ludwig M. Intranasal oxytocin: myths and delusions. *Biological Psychiatry* 2016 **79** 243–250. (https://doi.org/10.1016/j.biopsych.2015.05.003)

12 Winslow JT & Insel TR. Social status in pairs of male squirrel monkeys determines the behavioral response to central oxytocin administration. *Journal of Neuroscience* 1991 **11** 2032–2038. (https://doi.org/10.1523/JNEUROSCI.11-07-02032.1991)

13 Smith AS, Ågmo A, Birnie AK & French JA. Manipulation of the oxytocin system alters social behavior and attraction in pair-bonding primates, *Callithrix penicillata. Hormones and Behavior* 2010 **57** 255–262. (https://doi.org/10.1016/j.yhbeh.2009.12.004)

14 Dal Monte O, Noble PL, Turchi J, Cummins A & Averbeck BB. CSF and blood oxytocin concentration changes following intranasal delivery in macaque. *PLoS ONE* 2014 **9** e103677. (https://doi.org/10.1371/journal.pone.0103677)

15 Hughes AM, Everitt BJ, Lightman SL & Todd K. Oxytocin in the central nervous system and sexual behaviour in male rats. *Brain Research* 1987 **414** 133–137. (https://doi.org/10.1016/0006-8993(87)91333-3)

16 Insel TR, Young L & Wang Z. Central oxytocin and reproductive behaviours. *Reviews of Reproduction* 1997 **2** 28–37. (https://doi.org/10.1530/ror.0.0020028)

17 Calcagnoli F, de Boer SF, Althaus M, den Boer JA & Koolhaas JM. Antiaggressive activity of central oxytocin in male rats. *Psychopharmacology* 2013 **229** 639–651. (https://doi.org/10.1007/s00213-013-3124-7)

18 Guoynes CD, Simmons TC, Downing GM, Jacob S, Solomon M & Bales KL. Chronic intranasal oxytocin has dose-dependent effects on central oxytocin and vasopressin systems in prairie voles (*Microtus ochrogaster*). *Neuroscience* 2018 **369** 292–302. (https://doi.org/10.1016/j.neuroscience.2017.11.037)

19 McCarthy MM & Altemus M. Central nervous system actions of oxytocin and modulation of behavior in humans. *Molecular Medicine Today* 1997 **3** 269–275. (https://doi.org/10.1016/S1357-4310(97)01058-7)

20 Clark CL, St. John N, Pasca AM, Hyde SA, Hornbeak K, Abramova M, Feldman H, Parker KJ & Penn AA. Neonatal CSF oxytocin levels are associated with parent report of infant soothability and sociability. *Psychoneuroendocrinology* 2013 **38** 1208–1212. (https://doi.org/10.1016/j.psyneuen.2012.10.017)

21 Carson DS, Berquist SW, Trujillo TH, Garner JP, Hannah SL, Hyde SA, Sumiyoshi RD, Jackson LP, Moss JK, Strehlow MC, *et al*. Cerebrospinal fluid and plasma oxytocin concentrations are positively correlated and negatively predict anxiety in children. *Molecular Psychiatry* 2015 **20** 1085–1090. (https://doi.org/10.1038/mp.2014.132)

22 Horvat-Gordon M, Granger DA, Schwartz EB, Nelson VJ & Kivlighan KT. Oxytocin is not a valid biomarker when measured in saliva by immunoassay. *Physiology and Behavior* 2005 **84** 445–448. (https://doi.org/10.1016/j.physbeh.2005.01.007)

23 Szeto A, McCabe PM, Nation DA, Tabak BA, Rossetti MA, McCullough ME, Schneiderman N & Mendez AJ. Evaluation of enzyme immunoassay and radioimmunoassay methods for the measurement of plasma oxytocin. *Psychosomatic Medicine* 2011 **73** 393–400. (https://doi.org/10.1097/PSY.0b013e31821df0c2)

24 McCullough ME, Churchland PS & Mendez AJ. Problems with measuring peripheral oxytocin: can the data on oxytocin and human behavior be trusted? *Neuroscience and Biobehavioral Reviews* 2013 **37** 1485–1492. (https://doi.org/10.1016/j.neubiorev.2013.04.018)

25 Leng G & Sabatier N. Measuring oxytocin and vasopressin: bioassays, immunoassays and random numbers. *Journal of Neuroendocrinology* 2016 **28** [epub]. (https://doi.org/10.1111/jne.12413)

26 Crockford C, Wittig RM, Langergraber K, Ziegler TE, Zuberbuhler K & Deschner T. Urinary oxytocin and social bonding in related and unrelated wild chimpanzees. *Proceedings Biological Sciences* 2013 **280** 20122765. (https://doi.org/10.1098/rspb.2012.2765)

27 Samuni L, Preis A, Mundry R, Deschner T, Crockford C & Wittig RM. Oxytocin reactivity during intergroup conflict in wild chimpanzees. *PNAS* 2016 **114** 201616812. (https://doi.org/10.1073/pnas.1616812114)

28 Schaebs FS, Marshall-Pescini S, Range F & Deschner T. Analytical validation of an enzyme immunoassay for the measurement of urinary oxytocin in dogs and wolves. *General and Comparative Endocrinology* 2019 **281** 73–82. (https://doi.org/10.1016/j.ygcen.2019.05.015)

29 Schaebs FS, Wolf TE, Behringer V & Deschner T. Fecal thyroid hormones allow for the noninvasive monitoring of energy intake in capuchin monkeys. *Journal of Endocrinology* 2016 **231** 1–10. (https://doi.org/10.1530/JOE-16-0152)

30 Behringer V, Hohmann G, Stevens JMG, Weltring A & Deschner T. Adrenarche in bonobos (*Pan paniscus*): evidence from ontogenetic changes in urinary dehydroepiandrosterone-sulfate levels. *Journal of Endocrinology* 2012 **214** 55–65. (https://doi.org/10.1530/JOE-12-0103)

31 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017.

32 Field AP. *Discovering Statistics Using SPSS: (and Sex, Drugs and Rock 'n' Roll)*. Los Angeles: SAGE Publications, 2009.

33 Ziegler TE. Measuring peripheral oxytocin and vasopressin in nonhuman primates. *American Journal of Primatology* 2018 **80** e22871. (https://doi.org/10.1002/ajp.22871)

34 Wasser SK, Azkarate JC, Booth RK, Hayward L, Hunt K, Ayres K, Vynne C, Gobush K, Canales-Espinosa D & Rodríguez-Luna E. Non-invasive measurement of thyroid hormone in feces of a diverse array of avian and mammalian species. *General and Comparative Endocrinology* 2010 **168** 1–7. (https://doi.org/10.1016/j.ygcen.2010.04.004)

35 Behringer V, Deschner T, Murtagh R, Stevens JMG & Hohmann G. Age-related changes in thyroid hormone levels of bonobos and chimpanzees indicate heterochrony in development. *Journal of Human Evolution* 2014 **66** 83–88. (https://doi.org/10.1016/j.jhevol.2013.09.008)

36 Umapathy G, Kumar V, Wasimuddin KM, Kabra M & Shivaji S. Detection of pregnancy and fertility status in big cats using an enzyme immunoassay based on 5α-pregnan-3α-ol-20-one. *General*

and Comparative Endocrinology 2013 **180** 33–38. (https://doi.org/10.1016/j.ygcen.2012.10.009)

37 Bienboire-Frosini C, Chabaud C, Cozzi A, Codecasa E & Pageat P. Validation of a commercially available enzyme immunoassay for the determination of oxytocin in plasma samples from seven domestic animal species. *Frontiers in Neuroscience* 2017 **11** 524. (https://doi.org/10.3389/fnins.2017.00524)

38 Hawe A, Poole R, Romeijn S, Kasper P, van der Heijden R & Jiskoot W. Towards heat-stable oxytocin formulations: analysis of degradation kinetics and identification of degradation products. *Pharmaceutical Research* 2009 **26** 1679–1688. (https://doi.org/10.1007/s11095-009-9878-2)

39 Carter CS, Pournajafi-Nazarloo H, Kramer KM, Ziegler TE, White-Traut R, Bello D & Schwertz D. Oxytocin: behavioral associations and potential as a salivary biomarker. *Annals of the New York Academy of Sciences* 2007 **1098** 312–322. (https://doi.org/10.1196/annals.1384.006)

40 Amico JA, Ulbrecht JS & Robinson AG. Clearance studies of oxytocin in humans using radioimmunoassay measurements of the hormone in plasma and urine. *Journal of Clinical Endocrinology and Metabolism* 1987 **64** 340–345. (https://doi.org/10.1210/jcem-64-2-340)

41 Seltzer LJ & Ziegler TE. Non-invasive measurement of small peptides in the common marmoset (*Callithrix jacchus*): a radiolabeled clearance study and endogenous excretion under varying social conditions. *Hormones and Behavior* 2007 **51** 436–442. (https://doi.org/10.1016/j.yhbeh.2006.12.012)

42 Aroskar JP, Chan WY, Stouffer JE, Schneider CH, Murti VVS & Du Vigneaud V. Renal excretion and tissue distribution of radioactivity after administration of tritium-labeled oxytocin to rats. *Endocrinology* 1964 **74** 226–232. (https://doi.org/10.1210/endo-74-2-226)

43 Heistermann M. Non-invasive monitoring of endocrine status in laboratory primates: methods, guidelines and applications. *Advances in Science and Research* 2010 **5** 1–9. (https://doi.org/10.5194/asr-5-1-2010)

44 Behringer V & Deschner T. Non-invasive monitoring of physiological markers in primates. *Hormones and Behavior* 2017 **91** 3–18. (https://doi.org/10.1016/j.yhbeh.2017.02.001)

45 Robinson N, Saudan C, Sottas PE, Mangin P & Saugy M. Performance characteristics of two immunoassays for the measurement of urinary luteinizing hormone. *Journal of Pharmaceutical and Biomedical Analysis* 2007 **43** 270–276. (https://doi.org/10.1016/j.jpba.2006.06.034)

46 Heistermann M & Higham JP. Urinary neopterin, a non-invasive marker of mammalian cellular immune activation, is highly stable under field conditions. *Scientific Reports* 2015 **5** 16308. (https://doi.org/10.1038/srep16308)

47 Reyes TL, Galinsky AM, Hoffmann JN, You HM, Ziegler TE & McClintock MK. Social peptides: measuring urinary oxytocin and vasopressin in a home field study of older adults at risk for dehydration. *Journals of Gerontology: Series B, Psychological Sciences and Social Sciences* 2014 **69** (Supplement 2) S229–S237. (https://doi.org/10.1093/geronb/gbu104)

48 Preis A, Samuni L, Mielke A, Deschner T, Crockford C & Wittig RM. Urinary oxytocin levels in relation to post-conflict affiliations in wild male chimpanzees (Pan troglodytes verus). *Hormones and Behavior* 2018 **105** 28–40. (https://doi.org/10.1016/j.yhbeh.2018.07.009)

49 Moscovice LR, Surbeck M, Fruth B, Hohmann G, Jaeggi AV & Deschner T. The cooperative sex: sexual interactions among female bonobos are linked to increases in oxytocin, proximity and coalitions. *Hormones and Behavior* 2019 **116** 104581. (https://doi.org/10.1016/j.yhbeh.2019.104581)