university of groningen

University Medical Center Groningen

University of Groningen

Neglected aspects of hormone mediated maternal effects

Kumar, Neeraj

DOI:
10.33612/diss.101325389

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Kumar, N. (2019). *Neglected aspects of hormone mediated maternal effects: Studies on early embryonic modulation of maternal hormonal signals in avian eggs and related methodological aspects*. University of Groningen. https://doi.org/10.33612/diss.101325389

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 07-12-2022

Chapter 2

Substantial differences in testosterone measure between LC-MS/MS and radioimmunoassay (RIA) in bird and rat plasma, and in bird eggs, warrant caution for use of RIA kits

Neeraj Kumar, Martijn van Faassen, Bonnie de Vries, Josue Almansa, Bin-Yan Hsu, Asmoro Lelono, Ido Kema, Manfred Gahr, Ton G.G. Groothuis

ABSTRACT

Over the last two decades the study of hormone mediated maternal effects has flourished considerably, with using bird eggs as the predominant model. The most commonly used method for hormone measurements in bird eggs and calibrating hormone injections in ovo is by using commercially available radioimmunoassay (RIA) kits. However, RIA kit antibodies are characterized for usage in specific species' serum or plasma and not for egg yolk or even plasma of other species. Due to matrix effects, RIA may therefore give unreliable estimates when not accompanied by extensive purification steps and validated for specificity of the kit antibodies. Therefore, we compared concentrations of testosterone (T), one of the most extensively studied hormones, measured both by using a commercial RIA kit and by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) from the same samples, as the latter method is highly specific. We used egg yolk of three bird species, and blood plasma of two bird species and a rodent species. The results demonstrate that the commercial RIA kit after classical ether extraction gives much higher values than LC-MS/MS. The difference between both analytical technics differs according to species, tissue, and egg laying order, and increases with increasing T concentrations. As most studies use commercial RIA kits for calibrating experimental manipulation of steroids without extensive extraction and purification steps and antibody characterization, their outcome may reflect unnatural pharmacological effects, which warrants caution for the eco-evolutionary interpretations of hormone mediated maternal effects and for reliability of comparative studies.

1 INTRODUCTION

There is a strong interest in maternal steroid hormones as potential mediators of maternal effects in species across a wide array of taxonomic groups, including fish (Brown et al. 1988), reptiles (Radder 2007), birds (Gil 2008; Groothuis et al. 2005b; von Engelhardt & Groothuis 2011), as well as mammals (Braun et al. 2013; Del Giudice 2012). Bird species have been used as a particularly suitable model because their embryos develop outside the mother's body in the sealed egg environment, facilitating measurements and manipulation of maternal hormone deposition, while their ecology and evolution are well known (Groothuis et al. 2005b).

Manipulating steroids in bird egg has revealed a wide array of effects on the offspring's morphological, physiological, and behavioural traits (Gil 2008; von Engelhardt & Groothuis 2011). The dose of steroid manipulation is calculated based on the natural levels that are mostly measured using radioimmunoassays (RIAs). However, the antibodies used in RIA kits are characterized for a particular tissue from a particular species, i.e. a particular matrix. Their use in any other matrix might give unreliable measures due to uncharacterized antibody cross reactivity with often unknown matrix substances. One possibility to deal with this matrix problem is to remove naturally occurring steroids from the egg matrix by using charcoal and then add a known amount of a target steroid to characterize the RIA kit for that particular steroid in that particular matrix. This has only been done in a few cases (e.g. (Hahn 2011; Navara & Pinson 2010; Williams et al. 2005)) and, more importantly, does not really solve the issue. The use of charcoal or any other such substance could potentially not only remove the hormone but also part of the matrix, so that the validation of the antibody is then performed in a changed matrix that cannot be translated to the original matrix.

Tandem mass spectrometry in combination with liquid or gas chromatography (LC- or GC-MS/MS) is a much more reliable method as discussed extensively in the literature (Shackleton 2010; Taylor et al. 2015; Wudy et al. 2018). Instead of using antibodies, mass spectrometry is based on detecting mass to charge ratio of fragmented ions of a target compound. This method interpolates the concentrations of a target steroid from a calibration curve using the ratio of a known amount of added stable isotope labelled target compound (deuterated or preferably 13-Carbon standards) as an internal standard to the unlabelled target compound. As the labelled internal standard is structurally and chemically nearly identical to the unlabelled target steroid and is added before starting the procedures for hormone extractions, it corrects for matrix effects and accounts for recovery losses during the sample preparation procedure and for possible ion suppression in the mass spectrometer.

Though some studies have measured steroids in bird eggs by mass spectrometry (Chung & Lam 2015; De Baere et al. 2015; Hartmann et al. 1998; Larsen et al. 2015; Merrill et al. 2018; Mi et al. 2014; Sas et al. 2006; Tölgyesi et al. 2017; Wang et al. 2010; Xu et al. 2009; Yang et

al. 2008), the direct comparison of mass spectrometry data with RIA data to quantify the difference between the two methods in the same samples, for both within- and between-species differences, as well as between-tissue effects (such as plasma versus egg) is lacking. Likewise, in spite of a growing literature on the differences between immunoassays and mass spectrometry analyses for serum/plasma, mostly done in clinical diagnostics (Büttler et al. 2014; Dorgan et al. 2002; Fitzgerald & Herold 1996; Taieb et al. 2003; Wang et al. 2004), there is a lack of studies comparing the two methods for between-species and between-tissue effects. However, discrepancies between both methods may substantially influence the interpretation of the biological role of steroids, including the mechanism and function/evolution of hormone mediated maternal effects. We therefore compared the results from RIA and LC-MS/MS for the determination of T concentrations, as this hormone is one of the most commonly studied steroid. We initially focussed on bird eggs as, despite the very many papers on egg yolk testosterone, RIA's are characterized for plasma and not yolk. We used eggs of three bird species that differ in their diet: the black-headed gull (hereafter referred simply as gull) (*Chroicocephalus ridibundus*), the rock pigeon (*Columba livia*), and the red jungle fowl (*Gallus gallus*), and also compared eggs of the same mother that differ in laying order within the same nest. Two recent studies reported a lack of 5α-dihydrotestosterone (5α-DHT) in bird egg (Kumar et al. 2018a,b) when measured by mass spectrometry, even though RIA studies often report 5α-DHT in bird egg (e.g. (Elf & Fivizzani 2002; Schwabl 1997a)). Therefore, we also analysed egg yolk samples from all three bird species for the presence of 5α-DHT. In addition, we analysed blood plasma (hereafter simply referred as plasma) samples from three species: red jungle fowl (*Gallus gallus*), homing pigeon (*Columba livia domestica*), and rat (*Rattus rattus*), to assess between-species effects for plasma T measures, as well as to compare effects due to different matrices in egg yolk and plasma.

2 MATERIALS AND METHODS

2.1 Sample collection

To compare yolk T concentrations between results from RIA and mass spectrometry, we used yolk of freshly laid eggs collected from gull (16 eggs of 8 mothers), rock pigeon (12 eggs of 8 mothers), and red jungle fowl (12 eggs of 12 mothers) from the breeding stocks housed at the University of Groningen. The birds were housed in the outdoor aviaries (45m long x 9.6m wide x 3.75m high) under natural light and temperature conditions, and with *ad libitum* access to food and fresh water. Gulls were provided food pellets which contained fish content, whereas the other bird species are grain eaters that differ slightly in diet (see

food composition below). Eggs were stored at -20ºC until sample preparation, hormone extraction, and assays took place. In addition, we compared T concentrations measured by the two methods in the plasma from males (housed separate from females) of homing pigeon (n = 8, aged between 9-10 years), red jungle fowl (n = 27, aged 2 years), and wistar rat (n = 8, aged between 3-9 months). Plasma samples were also stored at -20ºC until hormone extraction and assays took place. Pigeons were fed a mixture of commercial pigeon seeds (Kasper Faunafood, article number 6721 and Kasper Faunafood, article number 6712), P40 vitamin supplement (Kasper Faunafood, article number P40), and small stones or grit. Red jungle fowls were fed daily laying pellets (Kasper Faunafood, article number 601820), once a week mixed grains (Kasper Faunafood, article number 384020). Gulls were fed daily Skretting E-3P Stella (Trouw Nutritions Nederland BV), three times a week 1.2 kg standard cat food (Arie Blok Nutritions, 3-mix, article number 6550) soaked in warm water. Rats were fed daily standard chow (Altromin International, article number Altromin 1414 Mod., 141005).

2.2 Sample preparation for hormone extraction

Yolks were separated from eggs. Yolk and plasma samples were thawed. Yolk samples were diluted with milliQ water and thoroughly mixed by vortexing. Samples were prepared for T extraction and concentration determination by RIA and LC-MS/MS using the same aliquots of yolk and plasma at the same time.

2.3 Hormone extraction and concentration determination by RIA

The hormone extraction for RIA, and RIA procedure itself were followed based on our previously published work (e.g. (Hsu et al. 2016)). Egg yolk and plasma samples were extracted using the following identical procedure. Each yolk (200 mg yolk-milliQ water homogenate in weight ratio of 1:1) and plasma (average 500 µl for pigeon, 407 µl for red jungle fowl, and 237 µl for rat) sample was added with 300 µl of milliQ water, and 50 µl of ^{3}H-labelled T (NET 553, Perkin Elmer) was added with 15 minutes of incubation at 37ºC to trace the recovery during extraction procedure. Each sample was extracted three times in 2.0 ml of a mixture of diethyl ether and petroleum benzene (volume ratio 7:3) by vortexing for 60 seconds (for 30 and 15 seconds respectively after second and third round of extraction), followed by centrifugation for 3 minutes at 2000 rpm at 4°C. The ether phase was decanted after snap freezing the tubes for 15 seconds in a mixture of ethanol-dry ice, and dried under a stream of nitrogen gas in the water bath at 37°C. The dried extract was resuspended in 2 ml of 70% methanol by vortexing and stored overnight at -20°C. Samples were centrifuged for 5 minutes at 2000 rpm at 4°C and the supernatant was dried under nitrogen gas in the water bath at 50°C.

The dried extracts were resuspended in phosphate buffered saline with gelatin (PBS-Gelatin buffer,5.30g $NaH_2PO_4.H_2O$; 16.35g $Na_2HPO_4.7H_2O$; 9.00 g NaCl; 1.00 g gelatin; 1.00 g NaN_3, adjusted to pH 7.10 using NaOH pellets) by thorough mixing on a vortex, using different amounts of the buffer depending on expected hormone concentrations; pigeon eggs: 250 µl for first eggs, 500 µl for last eggs (this species typically has a clutch size of two eggs per nest); red jungle fowl eggs: 450 µl; gull eggs: 2 ml for both first and last eggs (this species typically has a clutch size of three eggs per nest); pigeon plasma: 125 µl; red jungle fowl plasma: 125 µl; and rat plasma: 150 µl for samples from older rats and 200 µl from younger rats.

Of the extracts, 25 µl were used for T concentration determination using a commercial kit, TESTO-CT2 (Cisbio Bioassays), following the kit instructions. This kit uses anti-testosterone rabbit polyclonal antibodies bound to the solid phase (coated tubes), with very low cross reactivity with other steroids as per the information provided in the kit manual for human antiserum: 5α-dihydrotestosterone 2.6%, androstenedione 1.7%, methyltestosterone 0.3%, and other steroids < 0.1%. The assay is based on the principle of competitive binding of the added radiolabelled testosterone and the unlabelled testosterone contained in the calibrators provided in the kit (to plot the calibration curve) and the test samples (egg yolk or plasma) against a fixed and limited number of anti-testosterone antibody. The amount of radiolabelled testosterone bound to the antibody is inversely related to the amount of unlabelled testosterone present in the sample, and the latter can be interpolated using the calibration curve. Recoveries of the initially added radiolabelled T were measured in a subsample of this solution using scintillation cocktail (Ultima Gold, Perkin Elmer) and radioactivity counted on a scintillation counter. Standards were prepared using dilution series from pre-prepared stock and ranged from 0.08-20.00 ng/ml T. The average recoveries ranged from 86-88% for egg yolks (2.0-3.7% intra-assay coefficient of variation for RIA) and 65-86% for plasma (2.2-3.7% intra-assay coefficient of variation for RIA).

2.4 Hormone extraction and concentration determination by LC-MS/MS

For egg yolks, the hormone extraction for LC-MS/MS and the LC-MS/MS procedure itself were followed based on our previously published work (Kumar et al. 2018b,a). Each yolk sample (yolk-milliQ water homogenates in weight ratio of 1:1 for rock pigeon and red jungle fowl; 2:1 for gull) was added with an internal standard (25 µl of 30 nmol/L $^{13}C_3$ labelled T in 50% methanol, *IsoSciences*), thoroughly vortexed, and left for one hour at room temperature for equilibration. Since for each individual sample the ratio of the added internal standard to the target compound automatically corrected for any potential losses during extraction procedures as well as signal suppression due to differences in ionization efficiency in the mass spectrometer, hence the data were corrected for recovery losses. Each sample was extracted twice in 1 ml methanol by vortexing, followed by centrifugation

at 12000xg for 10 minutes at room temperature. The supernatant was transferred to tubes containing 200 mg of solid $ZnCl_2$ for lipid precipitation. The total volume of the combined supernatants was made to 4 ml by adding 2 ml methanol, and centrifuged at 12000xg for 10 minutes at 4°C. The supernatant was dried under nitrogen gas in a water bath at 50°C, re-suspended in 1 ml methanol, centrifuged at 12000xg for 10 minutes at room temperature, followed by addition of 1.8 ml water to the supernatant. This mixture was centrifuged at 12000xg for 10 minutes at 4°C. The supernatant was loaded on C18 SPE columns (3 ml, 500 mg, Grace Inc.) pre-equilibrated with 3 ml of methanol, followed by 3 ml of water. After collecting flow through, columns were washed with 3 ml water, and then eluted with 2 ml methanol. The eluent was dried under vacuum, re-suspended in 150 µl methanol, followed by addition of 350 µl water to make a final concentration of 30% methanol.

For plasma, 200 µL was added to a 2 mL 96-well polypropylene plate (Greiner Bio-One, Kremsmünster, Austria). To each well 25 µL of 30 nmol/L $^{13}C_3$ labelled T in 50% methanol (IsoSciences) internal standard working solution was added together with 25 µL pepsin solution (Labor Diagnostika Nord, Nordhorn, Germany). The samples were mixed by vortexing for 1 minute, and after incubation for 30 minutes at room temperature, ultrapure water was added to each well to a final volume of 1 mL. Subsequently, the plate was centrifuged (1500xg, 4°C, 30 minutes). Following centrifugation, the plate was placed in the autosampler.

All samples were analysed with a XEVO TQ-S tandem mass spectrometer (Waters Corp.), equipped with an Online SPE Manager and ACQUITY UPLC system (Waters Corp.). The UPLC flow rate was set at 0.4 ml/min using 10 mM ammonium acetate, 0.1% formic acid in water and methanol (containing 0.1% formic acid) as mobile phases A and B respectively. For each extract, 40 µl sample was injected for online SPE extraction on a XBridge C8 cartridge and chromatographic separation was performed on a Kinetex C18 column (2.1 x 100 mm, 2.6 µm). The mass spectrometer was operated under electrospray ionization mode with following operating conditions: cone voltage of 30 V, desolvation temperature of 600°C and source temperature of 150°C, and collision energy of 35 eV. Quantitative calibration was performed by using a calibration curve using the internal standard. The analysis was performed by monitoring two mass transitions. The monitored multiple reaction monitoring (MRM) transitions (m/z) were: 289 > 97 and 289 > 109 for T, 292 > 100 and 292 > 112 for $^{13}C_3$-T, 291 > 159 and 291 > 255 for DHT, 294 > 258 and 294 > 258 for $^{13}C_3$-DHT. The quantification limits were 0.025 nmol/L for T and 0.1 nmol/L for DHT. Inter-assay imprecision for human plasma was 2.8% at 0.37 nmol/L, 2.7% at 3.8 nmol/L, and 2.9% at 50 nmol/L, respectively. The assay was validated according to the Dutch guidelines for validation of analytical methods in medical laboratories by the Dutch Society of Clinical Chemistry and Laboratory Medicine and ISO15189 guidelines (Wielders et al. 2017).

2.5 Statistical analyses

First, it was determined whether the difference between the two methods (RIA values minus LC-MS/MS values) in the same samples of yolk and plasma was significantly different from zero by one-sample t-tests against zero value, after verifying the test assumptions. The effect sizes were calculated following Dunlap et al (Dunlap et al. 1996). For gulls, the averages of first and last laid eggs from each of the 8 individuals were used. For rock pigeons, out of the 8 individuals, both first and last laid eggs were available from 4 individuals and their averages were used, for the remaining 4 individuals only one egg per clutch was available and therefore data from individual eggs were used. For red jungle fowls, single eggs were sampled from 12 birds as previous work indicated no clear hormone difference between eggs over the laying sequence in contrast to the other bird species used in this study, in which second or third eggs have much more T than first eggs (Eising et al. 2001; Goerlich et al. 2009).

Next, we tested whether the difference between the two methods was larger in the lipid rich yolks than in plasma, using the data of those two species of which we had both egg and plasma samples (red jungle fowl and pigeon). This was analysed using a general linear model (GLM) with tissue as dependent variable, after verifying the test assumptions. We also tested to what extent species affect the difference between the two methods by analysing each tissue (egg yolk or plasma) separately, using a GLM with species as predictor (gull, pigeon, and red jungle fowl for egg data; pigeon, red jungle fowl, and rat for plasma data). In order to explore whether the differences between RIA and LC-MS/MS methods vary with concentration levels, (for example, larger differences with higher concentrations), which might confound the matrix effects in the other comparisons, the models included as covariate the T levels measured by LC-MS/MS (the most reliable method). Post-hoc tests were used for multiple comparisons with Bonferroni correction over marginal means.

Next, the within-species effect of egg laying sequence was determined in gulls for first (n=8) and last (n=8) laid eggs by controlling for nest (individual) identity as a random factor in a separate model, using T levels measured by LC-MS/MS as a covariate, after verifying the test assumptions.

Finally, the correlation coefficients were calculated for egg yolk samples from all three species, and separately for plasma samples from all three species, to have an overall estimate of the correlation between the two methods, independent of scale variation across species, for which reason standardized T concentrations (Z scores) were used after standardizing the data for each species and tissue separately.

3 RESULTS

Figure 1 shows a general trend of higher T concentrations when measured by RIA as compared to LC-MS/MS for egg yolk (panel A) and plasma (panel B) from three different species. Table 1 shows the difference between the two methods (RIA values minus LC-MS/MS values) for egg yolks and plasma samples, including effect sizes, from three different species, which were all significantly different from zero.

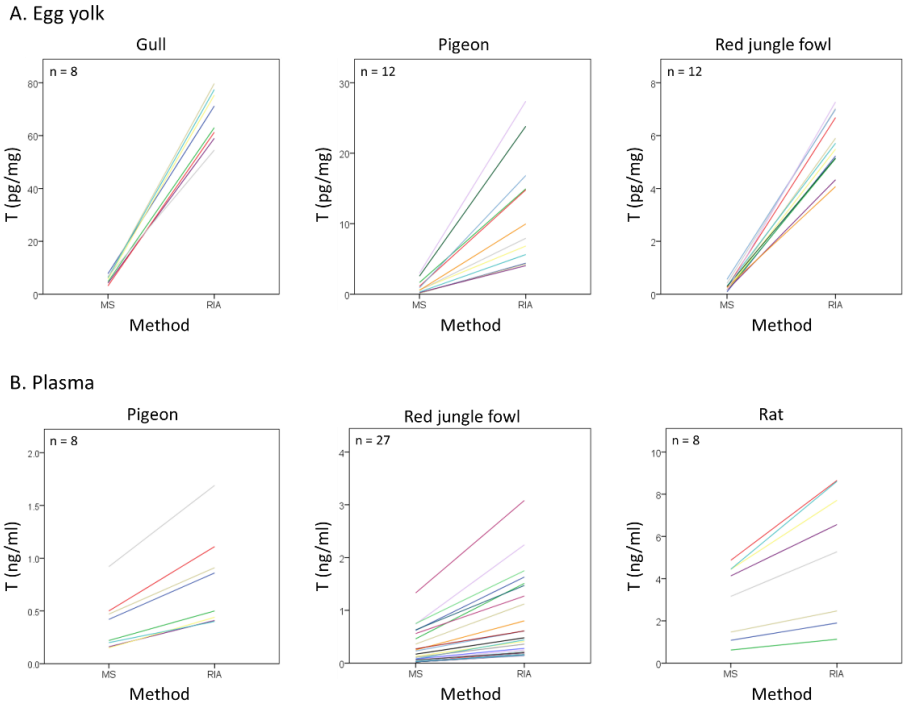

Figure 1. T concentrations measured by LC-MS/MS (labelled simply as MS) and RIA for (A) egg yolk and (B) plasma samples from three different species. Each line connects the samples from same individuals measured by both methods. Sample size is represented by n. Please note the differences in the scale of the y-axis, and that some of the lines overlap completely for egg yolks of pigeon and red jungle fowl, and plasma of red jungle fowl.

Table 1. One-sample t-tests depicting the difference between the two methods (RIA values minus LC-MS/MS values) being significantly different from zero for egg yolk, as well as for plasma, from each of the species.

Species	t	df	p-value	Mean difference (pg/mg for egg yolks, ng/ml for plasma)	95% confidence interval		Effect size
					lower	upper	
Egg yolks							
Gull	19.672	7	**< 0.001**	62.057	54.598	69.516	7.6
Pigeon	5.339	11	**< 0.001**	10.756	6.322	15.190	0.4
Red jungle fowl	19.015	11	**< 0.001**	5.500	4.863	6.137	5.7
Plasma							
Pigeon	5.879	7	**0.001**	0.410	0.245	0.575	0.3
Red jungle fowl	5.545	26	**< 0.001**	0.479	0.301	0.657	0.2
Rat	4.567	7	**0.003**	2.260	1.090	3.430	0.2

The GLM revealed that there was a significant effect of the tissue ($p < 0.001$), with the difference between the methods being much larger in egg yolk than in plasma (see Fig. 1), while the effect of the covariate itself was significant too ($p < 0.001$). When separate models were tested for egg yolks and plasma, the effect of species was significant for both egg yolks and plasma ($p < 0.001$), with a significant effect of the covariate for both egg yolks ($p = 0.001$) and plasma ($p < 0.001$) samples. The post-hoc tests for yolk showed that gulls differed significantly both from pigeons and red jungle fowls (Table 2A), the difference between the two methods being much larger for gulls (see Fig. 1), while pigeons did not differ significantly from red jungle fowls (Table 2A). The post-hoc tests for plasma showed that red jungle fowls did not differ significantly from pigeons, while rats differed significantly from red jungle fowls but not pigeons (Table 2B), the difference being larger for rats than the bird species.

There was a significant effect of the egg laying sequence ($p = 0.022$, Table 3), the difference between the two methods being much larger for the last laid eggs than the first.

The Pearson correlation coefficient (r) was 0.641 for egg yolks from all three species, and 0.988 for plasma from all three species, with the coefficient of determination (R^2) being 0.411 for egg yolks and 0.976 for plasma (Fig. 2).

5α-DHT was below the quantification limit (0.1 nmol/L) in egg yolks from all three bird species.

Table 2. Post-hoc tests (Bonferroni correction) for the effect of species on the difference between the two methods when egg yolks and plasma samples are analysed in separate models, using T levels measured by LC-MS/MS as a covariate.

(A) Egg yolks

Species compared	Mean difference	Standard error	p-value	Standardized coefficient (beta)
Gull – Pigeon	34.3	5.3	**< 0.001**	1.8
Gull – Red jungle fowl	36.9	6.0	**< 0.001**	1.9
Pigeon – Red jungle fowl	2.6	2.3	0.765	0.1

(B) Plasma

Species compared	Mean difference	Standard error	p-value	Standardized coefficient (beta)
Red jungle fowl – Pigeon	0.17	0.1	0.289	0.01
Pigeon – Rat	0.4	0.2	0.121	0.02
Red jungle fowl – Rat	0.5	0.2	**0.005**	0.03

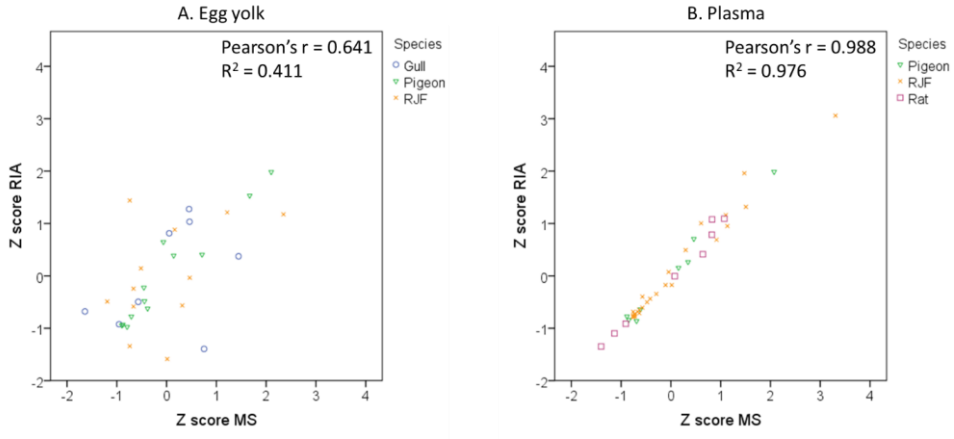

Figure 2. Scatterplot showing correlation between standardized (Z scores) T concentrations for the two methods – RIA and LC-MS/MS (labelled simply as MS), for (A) egg yolk and (B) plasma samples. The abbreviation RJF stands for red jungle fowl.

Table 3. The within-species effect of the egg laying sequence on the difference between the two methods in gull eggs, using T levels measured by LC-MS/MS as a covariate.

Egg laying order compared	Mean difference	Standard error	p-value	Standardized coefficient (beta)
Last – first	19.1	7.3	**0.022**	1.4

4 DISCUSSION

The results show that the RIA kit that we used, after applying very commonly used extraction procedures that typically do not include chromatographic purification steps such as solid phase extraction, measured substantially higher T concentrations as compared to LC-MS/MS and that this difference increased with higher T concentrations. The difference in T concentrations between the two methods cannot be attributed to the differences in recoveries during extraction procedures because T concentration values were corrected for recovery losses for each sample and for both methods. Another reason for higher RIA values could be due to use of tritium labelled T compared to $^{13}C_3$ labelled T in LC-MS/MS, where tritium label being relatively less stable could detach from T molecules, leading to underestimation of recoveries, and thus overestimation of corrected T values in RIAs. However, this seems not be the case as the recoveries for RIAs (reported above) were above 80% in most cases. The differences between the methods, even when corrected for T concentrations as measured with LC-MS/MS, indicate strong matrix effects, with a possible cross reactivity of the RIA antibodies with other steroids and/or other matrix substances. Interestingly, the difference between the methods was largest in gull eggs compared to pigeon and red jungle fowl eggs (Table 2A). The latter two species are seed eaters, whereas gulls were fed with trout pellets, containing much more fat, that is known to disturb RIA assays (Von Engelhardt & Groothuis 2005). Similarly, the difference between the methods is larger in yolk than in plasma (Fig. 1, Table 2), with the former containing much more fat. Apparently, fat is not sufficiently removed by the commonly used extraction methods. Therefore, it might be advantageous not only to include chromatographic purification but also additional lipid precipitation step such as using $ZnCl_2$, while using RIAs. The coefficient of determination (R^2) was much lower for egg yolks (0.411) than for plasma (0.976). Interestingly, the difference between the two methods is also much larger in last eggs than first eggs while controlling for baseline T concentrations and nest identity (Table 3).

5α-DHT, the most potent androgen (Fang 2003), is often reported to be present in the egg yolk in the studies based on RIAs (e.g. (Elf & Fivizzani 2002; Schwabl 1997a)). Intriguingly, 5α-DHT was not detectable in the egg yolk of any of the three bird species when analysed by LC-MS/MS, in spite of its quantification limit being as low as 0.1 nmol/L, warranting another caution for the reliability of classical radioimmunoassays for hormone analyses of

eggs. Although we did not test the recoveries for DHT, a previous study showed a similar level of recoveries for T and DHT when extracted in methanol (Upreti et al. 2015).

Our findings have several important implications. First, hormone manipulation based on their concentrations or total amount in the yolk or plasma as measured by commonly used RIAs, without extensive extraction and purification as well as antibody characterization, leads to a pharmacological dose and thus makes the biological interpretations of these manipulations unreliable. This is a serious issue as, since the discovery of T in bird eggs (Schwabl 1993), many hormone manipulations have been done in ovo based on RIA's, leading to an extensive body of literature (e.g. reviewed in (Gil 2008; von Engelhardt & Groothuis 2011). Likewise, even many more studies have been conducted in which T concentrations in plasma were manipulated by injections or hormone releasing implantations, to analyse its effect on a wide array of traits (Adkins-Regan 2005; Nelson 2011). Our results warrant reinterpretation of such studies, especially when interpreting the results in a functional or evolutionary framework. One could argue that pharmacological doses may still induce normal effects due to a ceiling effect when all receptors are occupied by the hormone, but dose dependent effects have been demonstrated in many cases (Muriel et al. 2015; Podmokła et al. 2018; von Engelhardt & Groothuis 2011). Second, as the over-estimation by commonly used invalidated RIAs depends on the species as well as baseline levels of the hormone of interest, this makes between-species comparisons, including both ecological and evolutionary approaches and meta analyses, unreliable. Third, even when working with the same species, the effect of egg quality as indicated by the laying order effect, and the effect of base levels for yolk can still lead to serious misinterpretations when using invalidated RIA.

We used a RIA kit designed for human serum. As RIA kits are mostly validated for serum/plasma samples, it might often be assumed that they are suitable to use for plasma samples from any species and not even standard extraction protocols are always applied. Intriguingly, we show that RIAs give higher estimates as compared to LC-MS/MS even for plasma samples of birds as well as rats when measured by using a kit for human serum, partly due to baseline differences in the hormone levels between different species and partly due to other matrix effects. Perhaps the difference between the two methods would have been smaller if we would have used separate RIA kits with their antibodies designed for plasma of the rat and of the bird species. Therefore, RIA kits, including their antibodies, should be tested and optimized for species- and tissue-specific use. For plasma, the effect sizes for the difference between the two methods itself (Table 1) as well as for the effect of species (Table 2) are small, and the coefficient of determination is large (Fig. 2). Nevertheless, even after applying a standard extraction procedure, which is not even always done in case of plasma hormone concentration determination using RIAs, RIAs give substantially larger estimates, especially with higher hormone concentrations (Fig. 1).

In conclusion, we report that the use of RIAs, when not cautiously validated, is unreliable for experimental manipulation of hormones in plasma and especially in the egg, for interpretations of variations in maternal hormone allocation and the natural relation between plasma or yolk T and other traits, as well as between-species comparisons. Such effects shall be verified by using more reliable LC-MS/MS methods.

ETHICS

All the animal handling was conducted according to the established guidelines and regulations of the animal welfare committee of the University of Groningen, and all relevant procedures were approved by the committee under licenses 6835, 5635, 6710, and 5095.

ACKNOWLEDGEMENTS

We thank the animal caretakers – Saskia Helder, Diane ten Have, and Martijn Salomons. We also thank Kevin Matson and Jan Bruggink for their kind help with parts of sample collection.