Title: High-resolution mass spectrometry as an alternative detection method to tandem mass spectrometry for the analysis of endogenous steroids in serum.

Authors: Ponzetto F, Boccard J, Baume N, Kuuranne T, Rudaz S, Saugy M, Nicoli R

Journal: Journal of chromatography. B, Analytical technologies in the biomedical and life sciences

Year: 2017 May 1

Issue: 1052

Pages: 34-42

DOI: 10.1016/j.jchromb.2017.03.016

High-resolution mass spectrometry as an alternative detection method to tandem mass spectrometry for the analysis of endogenous steroids in serum

Federico Ponzetto[1], Julien Boccard[2,3], Norbert Baume[1], Tiia Kuuranne[1], Serge Rudaz[2,3], Martial Saugy[4], Raul Nicoli[1*]

[1]Swiss Laboratory for Doping Analyses, University Center of Legal Medicine Geneva and Lausanne, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Switzerland

[2]School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Switzerland

[3]Swiss Center of Applied Human Toxicology (SCAHT), University of Basel, Switzerland

[4]Center of Research & Expertise in anti-doping Sciences, University of Lausanne, Switzerland

*Corresponding author:

Dr. Raul Nicoli, Swiss Laboratory for Doping Analyses, Chemin des Croisettes 22,

1066 Epalinges, Switzerland, Tel.: +4121 314 73 30; E-mail address: raul.nicoli@chuv.ch

Abstract

Recently, steroid hormones quantification in blood showed a promising ability to detect testosterone doping and interesting complementarities with the urinary module of the Athlete Biological Passport (ABP). In this work, an ultra-high pressure liquid chromatography–high-resolution mass spectrometry (UHPLC-HRMS) method was developed for the quantification of eleven endogenous steroids in serum. The performance of the full scan and targeted SIM acquisition modes was evaluated and compared to the performance of tandem mass spectrometry (MS/MS). Passing-Bablok regressions and Bland-Altman plots were assessed for each analyte of interest, and concentration values measured by HRMS showed high correlation with the ones obtained by MS/MS for all target hormones, with low absolute differences in the majority of cases. A slight decrease in terms of sensitivity was observed with HRMS in both acquisition modes, but performing an analysis of variance multiblock orthogonal partial least squares (AMOPLS) on the dataset obtained with all three methods revealed that only 0.8% of the total variance was related to instrumentation and acquisition methods. Moreover, the evaluation of the testosterone administration effect over time highlighted testosterone itself and dihydrotestosterone as the most promising biomarkers of exogenous testosterone administration. This conclusion suggests that HRMS could provide suitable performance for blood steroid analysis in the anti-doping field.

Keywords: testosterone, doping, serum, UHPLC-HRMS, AMOPLS

1. Introduction

The detection of testosterone (T) abuse in sport is currently achieved through the steroidal module of the Athlete Biological Passport (ABP). The so-called "steroid profile" was implemented in 2014, and it is obtained from urinary concentrations of six endogenous anabolic androgenic steroids (EAAS) measured by gas chromatography-(tandem) mass spectrometry (GC-MS(/MS)). It consists of longitudinal monitoring, with individual reference intervals calculated by a Bayesian adaptive model, of five ratios known to be potentially altered following the administration of synthetic forms of EAAS: androsterone/testosterone (A/T), androsterone/etiocholanolone (A/Etio), 5α-androstane-3α,17β-diol/5β-androstane-3α,17β-diol (5αAdiol/5βAdiol), 5α-androstane-3α,17β-diol/Epitestosterone (5αAdiol/E) and the most important, testosterone/epitestosterone (T/E) [1]. Although this new approach clearly improved detection capabilities in comparison with the previous population threshold criterion of T/E > 4, it still suffers from major drawbacks. These include the large presence in the urine matrix of both endogenous (e.g., metabolism, ethnicity) and exogenous (e.g., ethanol, bacterial contamination) confounding factors, which can influence either the quantification of the urinary steroid profile or its interpretation [2], and the difficulty of detecting doping in individuals with a genetic polymorphism for UGT2B17 enzyme activity [3-6] and/or in athletes who have been administered T gel and patch formulations [7].

In the last few years, several studies have been conducted with the aim of improving steroid abuse detection capabilities. At first, major efforts were focused on the identification, by means of GC-MS(/MS) platforms, of new urinary biomarkers of EAAS abuse to be added to the panel of ratios included in the steroid profile [8-11]. Then, research broadened and oriented to UHPLC-

MS/MS methods as well, as this type of technique is faster, more sensitive and capable of detecting phase II metabolites of steroid hormones in their intact form [12-14], unlike GC-MS analysis which always requires a derivatization step, eventually preceded by a hydrolysis step. More recently, high-resolution mass spectrometry (HRMS) has also been used to provide innovation in the context of steroid analysis. Indeed, holistic approaches such as metabolomics (defined as steroidomics [15]), allowed the untargeted evaluation of a large number of compounds, thus also representing a promising strategy for the discovery of new biomarkers and metabolites for anti-doping purposes [16, 17]. Nevertheless, the use of HRMS platforms by anti-doping laboratories for screening analyses is constantly increasing [18-21], even if their use for quantification purposes is still not widespread.

In the anti-doping field, both endogenous and exogenous steroids have been traditionally analyzed in urine, but a first attempt of endogenous steroid analysis in blood was recently conducted [22]. The blood matrix is clearly more difficult to manipulate/contaminate than urine, and quantitative analysis of target compounds in blood represents a real snapshot of athlete physiological status, a key aspect for possibly better discriminating doping from pathologies. In addition, the study of steroid metabolism in blood could also be particularly relevant to explain diseases associated with a possible malfunction in steroidogenesis, in particular steroid metabolism enzymes (e.g. congenital adrenal hyperplasia, CAH) as well as in environmental sciences to better understand the mode of action of endocrine disrupting chemicals. On the other hand, the blood matrix also raises some concerns in the context of sports drug testing: its sampling is more invasive compared to urine, the collected volume is significantly lower and more accurate conditions for transportation and storage are needed.

In this work, a UHPLC-HRMS method was developed for the quantification of 11 endogenous steroid hormones in serum, including major androgens, progestogens and corticoids. The performance of the full scan (FS) and targeted selected ion monitoring (t-SIM) acquisition modes was evaluated and compared to that of UHPLC-MS/MS obtained in a previous work using samples from a testosterone clinical study [22]. Furthermore, in addition to quantitative performance, the ability to describe the exogenous testosterone administration by the three acquisition methods was also evaluated by means of an ANOVA-based multivariate statistical analysis.

2 Materials and methods

2.1 Study samples

Serum samples used for the method comparison were obtained from a T administration clinical trial (Swissmedic protocol n° 155/11) with 19 healthy male volunteers, who were administered twice with a T transdermal patch (Testopatch® 2.4 mg/24 h, Pierre Fabre Pharma GMBH, Freiburg, Germany) followed by T undecanoate capsules (Andriol Testocaps®, Essex Chemie AG, Luzern, Switzerland). The detailed description of the study is presented elsewhere [7].

2.2 Chemicals and reagents

Available standards of endogenous steroids were purchased from Lipomed (Arlesheim, Switzerland), Cerilliant (Round Rock, TX, USA), Steraloids (Newport, RI, USA) and Sigma-Aldrich (Buchs, Switzerland), while all labeled internal standards (IS) were provided by the Australian Government National Measurement Institute (Pymble, Australia). Methanol (MeOH) was purchased from Macron Fine Chemicals (Deventer, Netherlands), and acetonitrile (ACN) ULC/MS (> 99%) and formic acid (FA) ULC/MS (99%) were supplied by Biosolve BV (Valkenswaard, Netherlands). Charcoal-dextran stripped human serum was obtained from Dunn Labortechnik GmbH (Asbach, Germany). Deionized water was obtained by a Milli-Q®-grade system (Millipore, USA) and was used for the preparation of all buffers and solutions.

Analytes and IS mixture solutions were prepared in MeOH at appropriate concentrations and used for the preparation of calibration samples in depleted serum. Linear calibration curves were created for each analyte (weighting 1/x) to measure steroid concentration, and calibration samples were used for the extraction of lower limits of quantification (LLOQ) and accuracy

values of the method. Details concerning the composition and concentration of the calibration samples and IS mixtures are given in Supplementary Material Tables S1 and S2.

2.3 Sample preparation and UHPLC conditions

All the details concerning the extraction protocol and chromatographic conditions have been described in [22]. Briefly, a supported liquid extraction SLE+ 400 µL (Biotage, Uppsala, Sweden) in a 96 well-plate format was used to extract steroid hormones. Each serum sample (200 µL) was spiked with 20 µL of the IS mixture, diluted with 200 µL of water and then agitated for few minutes. Then, 400 µL of sample were loaded on each well and positive pressure of 3 psi was applied for 30 seconds to facilitate sample adsorption; after 5 minutes, the elution was carried out by adding 700 µL of DCM to each well and applying a pressure of 3 psi for 1 minute. After evaporation, 10 µL of reconstituted extracts were injected on a UPLC chromatographic system (Waters, Milford, MA, USA) equipped with an Ethylene Bridged Hybrid (BEH) C_{18} column (100 x 2.1 mm, 1.7 µm; Waters) set at 30°C and a pre-column. The mobile phases consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in ACN, and the flow rate was set at 400 µL/min. The gradient started linearly from 2% to 25% B over 0.5 min, followed by an increase to 58% B over 5.5 min and by a further increase to 98% B over 2 min; the column was then re-equilibrated for 3 min at initial conditions.

2.4 MS conditions

2.4.1 MS/MS analysis

The UPLC system was coupled to a Xevo-TQ S triple quadrupole MS/MS system from Waters working in ESI positive and selected reaction monitoring (SRM) mode. The detailed instrumental UHPLC-MS/MS conditions (SRM transitions, ESI conditions, cone voltages and collision energies) are described in a previous article [22].

2.4.2 HRMS analysis

HRMS analyses were performed by coupling the UPLC system to a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Mass calibration (<3 ppm) was performed once a week using the Pierce® LTQ Velos ESI Positive Ion Calibration standard mixture (Thermo Fisher Scientific) containing *n*-butylamine, caffeine, MRFA (peptide of Met-Arg-Phe-Ala acetate salt) and Ultramark 1621. Detection of the targeted steroids was performed in positive ESI in both FS and t-SIM acquisition modes. The heated ESI source (HESI II) was used with a probe heater temperature of 425°C, and the sheath gas and auxiliary gas pressures were set to 50 and 15 arbitrary units, respectively. The sweep gas flow was set to 3 arbitrary units. The ion spray voltage was set to 4.5 kV, the capillary temperature to 250°C and the S-Lens RF level was 55%. FS mass spectra were acquired using a mass resolution of 70,000 (full width at half maximum, FWHM) at m/z 200, with a maximum IT fill time of 250 ms and the automatic gain control (AGC target) set to $1e^6$. The acquired mass range was from m/z 200 to 600. The t-SIM acquisition mode was also performed using a mass resolution of 70,000 FWHM, with a maximum IT fill time of 250 ms and the AGC target set to $5e^4$. An isolation window of 0.4 m/z was set, and the maximum number of precursor ions to be multiplexed in a scan event (MSX

count) was from 2 to 8, depending on the proximity of the analytes in terms of retention times. All extracted chromatograms were obtained using an extraction window of 10 ppm.

2.5 Comparison of quantitative performance

The comparison between HRMS and MS/MS methods was performed employing two different approaches. First, the correlation of the steroid concentrations obtained with the two methods was evaluated using the Passing-Bablok regression together with Bland-Altman plots; the latter helped evaluate the bias, as they correspond to regression residual plots. Then, a multivariate data analysis, involving analysis of variance multiblock orthogonal partial least squares (AMOPLS), was applied to simultaneously investigate signal variations related to the different experimental factors, i.e., inter-individual, temporal and method-related sources of variability, and their potential interactions. Indeed, AMOPLS allows the reliability of experimental effects to be objectively evaluated using a specific effect-to-residual ratio. The aim of this second approach was not to provide an absolute evaluation of the variations caused by the different quantification methods but to fairly compare and establish the relative impact of each source of variability in the obtained dataset. Random permutations of the design matrix simulate data under the null hypothesis (i.e., no effect), and the values obtained can be compared to the experimental value. AMOPLS models were computed under the MATLAB 8 environment (The MathWorks, Natick, MA, USA) and compared to a series of 10^3 random permutations.

3. Results and discussion

3.1 Optimization of HRMS conditions

The chromatographic conditions were already optimized in a previous study [22] allowing a satisfactory separation of the analytes, especially in the case of 11-deoxycortisol and corticosterone, which had the same exact mass and close retention times.In this work, the method development was mainly focused on the optimization of HRMS conditions. Details of the detection parameters are presented in Table 1 for the target analytes and in Supplementary Material Table S3 for the labeled internal standards (IS). First, all the ESI source parameters were tuned to obtain the best sensitivity. As DHT proved to have the lowest ionization efficiency among the 11 analytes, it was chosen as the target compound, and direct infusion of a DHT standard in MeOH at 1 µg/mL was performed. In addition, this compound had already been highlighted as a promising biomarker of testosterone administration in serum [22].

Then, the first investigated acquisition mode was the FS. The scan range was set from 200 to 600 m/z, allowing the detection of all steroid hormone ions and their potential adducts; the resolution, AGC target and Maximum IT were optimized with the aim of obtaining a sufficient number of acquisition data points over each chromatographic peak. For that, a target range of 10-15 data points was set for acquisition. An example chromatogram for 17αOH-progesterone at 1 ng/mL obtained in FS mode is shown in Figure 1A, where 16 data points over the peak were obtained with a resolution of 70,000 FWHM. Because some of the target steroids, such as epitestosterone, progesterone and DHT, are known to be at a very low concentration in serum, a t-SIM experiment was also evaluated with the aim of enhancing sensitivity. For this purpose, preliminary experiments using a single t-SIM event were carried out during the whole

chromatographic run. For this experiment, an inclusion list containing m/z values of all target hormones and relative internal standards ions was created, and the acquisition was performed setting the multiplexing (MSX) count to 1. As shown in Figure 1B, this strategy was not satisfactory to obtain a sufficient number of data points across the peak, also compromising sensitivity. Indeed, in the retention time window where 17αOH-progesterone eluted (from 5 to 5.5 min), three other analytes, as well as their respective IS, were detected, reaching a total of 8 compounds (see Table 1). This meant that in this specific time window, after each FS event, only one of the 8 present ions was allowed to enter the C-Trap and be analyzed, resulting in a limited number of acquisitions in t-SIM mode for all the compounds in this time period (7 points under the peak in the case of Figure 1B). To resolve this issue, a different t-SIM experiment was then performed making use of the MSX function, depending on the number of compounds co-eluting in the same time window. MSX values from 2 (only the target analyte and its internal standard) up to 8 (four analytes and their relative IS) were selected (see Table 1). After each FS event, the ions of all co-eluting multiplexed compounds were accumulated in the C-trap and subsequently analyzed. This second strategy proved to be efficient, allowing a number of acquisition points similar or equal to the FS acquisition for all target steroids. As observed in Figure 1C, by applying an MSX count of 8, it was possible to obtain a very satisfactory peak shape with approximately 16 data points for 17αOH-progesterone with the multiplexed t-SIM acquisition mode.

Concerning sensitivity, the t-SIM acquisition provided an increased signal to noise (S/N) ratio, in particular for DHT, one of the two blood markers of T administration previously highlighted: a significant increase in the S/N from 9 to 32 RMS (root mean square) was observed passing from FS to multiplexed t-SIM, despite no remarkable augmentation of peak area. For analytes with a

low concentration and/or poor ionization in the ESI source, the t-SIM acquisition mode could, therefore, represent the optimal solution for quantification purposes, assuming the utilization of MSX count.

3.2 Comparison of quantitative performance

Once the HRMS conditions were optimized, 491 serum samples collected during a testosterone administration clinical study were prepared and analyzed using both of the developed UHPLC-HRMS and UHPLC-MS/MS methods. For comparison purposes, concentrations of 11 target steroids were measured by means of the three different acquisition modes: FS, multiplexed t-SIM and SRM.

Epitestosterone, progesterone and deoxycorticosterone were present in serum at very low concentrations (below 100 pg/mL), and it was not possible to detect them with the HRMS platform; therefore, they were discarded for the following comparison.

3.2.1 Assay correlation

For each of the 8 sufficiently concentrated compounds, Passing-Bablok regressions [23] and Bland-Altman plots [24, 25] were assessed, using MS/MS values as a reference method. The plots obtained for T and DHT are presented in Figures 2 and 3, respectively, and discussed herein in more detail, while the graphs for all remaining compounds are shown in Supplementary Material Figures S1-S6.

For the Passing-Bablok regression, the parameters that were evaluated were the slope, the intercept, and the determination coefficient (R^2). For the Bland-Altman plots, the percentage of samples with a difference of measured concentrations between ±20% (arbitrary acceptability

range [26]) was evaluated. In both cases, the T concentration measured with the HRMS method correlated well with MS/MS. In the t-SIM vs SRM plot (Figure 2A), the samples were less spread than in the FS vs SRM plot (Figure 2B), resulting in a better determination coefficient; this dispersion of the samples is more pronounced in the high concentration region. Moreover, a slope value close to one was found for both regressions, while the obtained intercept had a negative value in both cases, suggesting a minor underestimation of T concentration with both HRMS methods. According to the Bland-Altman plots, for a high percentage of samples (99.6% and 93.9%, respectively), measured T concentration differences were within the acceptability limits arbitrarily defined for both t-SIM (Figure 2C) and FS (Figure 2D); mean biases of -3.615% and 0.298% were observed comparing MS/MS measurements with the t-SIM and FS results, respectively.

In Figure 3, the plots of the DHT measurement comparison are presented. Looking at Passing-Bablok regressions, similar results to T were obtained, demonstrating a good correlation between different quantification results, considering slope values close to one and satisfactory R^2 for both HRMS methods (0.9328 t-SIM, 0.9013 FS). On the other hand, there was a lower percentage of samples within the acceptability limits (82.7% t-SIM, 73.5% FS) and a higher mean bias from SRM quantification (6.48% t-SIM, 6.89% FS) evaluating Bland-Altman plots. For concentrated analytes such as T, HRMS measurements appear to be a valuable alternative to classical MS/MS experiments. However, the higher dispersion of samples in Bland-Altman plots in the low concentration region suggests that HRMS measurements are slightly less sensitive than MS/MS. This could be an issue for the less concentrated analytes such as DHT, especially

when FS acquisition is considered. In this context, the t-SIM acquisition mode could reduce the gap with MS/MS quantification, thanks to a significant increase in sensitivity.

To further investigate this aspect, calibration curves were constructed and LLOQ values were estimated for all steroid hormones with both the FS and t-SIM developed methods. First, several regression models were tested to select the most suitable and simple response function. Because the hypothesis of variance homogeneity was rejected, the simplest regression model without data transformation, based on the least squares method, could not be retained. Therefore, the best weighting factor was chosen taking into account the relationship between the logarithm of the natural variance and the concentration as described elsewhere [27]. For each analyte, the specially selected weighting factor was the inverse of the concentration raised to the λth power ($1/x^{\lambda}$), λ being the slope of the line fitted to the data on the logarithm scale. Hence, the mean selected weighting factor was round off the inferior unit and determined to be $1/x$. Then, the lowest concentration at which the back-calculated concentration of the calibration curves exhibited a precision that did not exceed 20% of the CV and accuracy within 20% of the nominal concentration was evaluated. This approach, in accordance with official guidelines such as the FDA [28], permitted obtaining the LLOQ values shown in Table 2 together with those of the reference MS/MS method. The accuracy and repeatability details for all target analytes at all concentration levels are presented in Supplementary Material Tables S4 and S5.

3.2.2 Multivariate data analysis

In addition to the classical comparison of the analytical methods for each of the measured target compounds discussed above, a more global approach has also been applied to evaluate the quality of the measurements obtained using HRMS. A multivariate data analysis method based on ANOVA and PLS, namely AMOPLS, which was recently proposed for the analysis of omics data generated from designed experiments [29], was used for proper data analysis accounting for both the highly multivariate structure of the data and the study design. Indeed, multifactorial experiments, such as the analysis of clinical study samples, generate data simultaneously altered by several sources of variation. The systematic strategy proposed herein is particularly well suited to gather information on the potential effects of experimental factors and their interactions, and it was demonstrated to be a relevant tool to disentangle the influences of specific factors or interactions in multifactorial experiments, with a simplified interpretation of signal variations based on specific sets of scores and loadings related to each effect.

Two separate AMOPLS models were therefore computed based on the combination of quantitative data from the 8 steroid hormones measured by all three analytical methods in blood control samples and samples collected after transdermal and oral T administrations. Unit variance scaling was applied to analyte concentration variations, avoiding their impact in intensity range. The main effects of the three experimental factors taken into account for this study, namely inter-individual variability (*Volunteer* factor, 19 levels), longitudinal signal variations (*Time* factor, 16 levels) and differences due to the quantification strategy (*Method*

301 factor, 3 levels), as well as their first-order interactions (*Volunteer*Time*, *Volunteer*Method* and
302 *Time*Method*) were included in the models.

303 Concerning the transdermal administration, as a first investigation of signal variations, the
304 relative contribution related to each effect and interactions (expressed as sum of squares) was
305 computed to obtain the following values: *Volunteer* 29.7%, *Time* 38.6%, *Method* 0.8%,
306 *Volunteer*Time* 28.9%, *Volunteer*Method* 0.8%, *Time*Method* 0.2% and Residuals 1.0%. These
307 results, shown in Figure 4A, clearly highlight the strong influence of the inter-individual
308 variability in the measured steroids levels, together with marked alterations during the
309 temporal follow-up. Interestingly, the cumulative proportion of signal variations due to the
310 different quantification strategies (SRM and HRMS in both FS and t-SIM modes) was very low
311 (main effect+interactions <2%).
312 Random permutations of the experimental design were then carried out to evaluate the
313 statistical significance of the results. The comparison with random designs highlighted the main
314 effect due to the *Time* factor as the only significant structured source of variation in the dataset
315 ($p<0.01$). Notably, despite an important proportion of total variability related to the *Volunteer*
316 factor, all the other ANOVA terms were declared non-significant ($p>0.05$). This may be because
317 inter-individual variations are related to biological noise. While no major structure emerges
318 from these uncoordinated differences, these results also confirmed that the very low
319 contribution related to the *Method* factor was statistically negligible. Total inter-individual
320 variability is obtained by summing the main effect of the *Volunteer* factor (differences between
321 individuals on average over all time points) and the interaction effect between *Volunteer* and
322 *Time* (differences between the evolutions over time).

Predictive components related to the *Time* factor were then further investigated to highlight trends in the sample distributions (Supplementary Material Figure S7A). The most important temporal pattern was associated with a circadian modulation of steroid levels, clearly visible on the first latent variable associated with the *Time* factor (47.9%). This pattern confirms the results previously obtained by MS/MS and was mainly associated with androstenedione, DHEA, corticosterone, cortisol, and 11-deoxycortisol. Additionally, the effect of T transdermal intake was observable on the second component associated with the *Time* factor (30.3%) (Figure 4B), characterized by a marked increase until 12 h after administration and a slow decrease until 48 h. The major biomarkers associated with this trend were T and DHT.

In the same fashion as transdermal administration, differences due to inter-individual and temporal variability were responsible for the largest variations of the dataset of the oral testosterone undecanoate (TU) intake, as illustrated by the relative sum of squares related to each effect in Figure 5A: *Volunteer* 26.5%, *Time* 31.3%, *Method* 0.6%, *Volunteer*Time* 35.7%, *Volunteer*Method* 0.9%, *Time*Method* 0.4% and *Residuals* 4.5%. According to random permutations, the only significant effect was related to the *Time* factor ($p<0.01$), while all other effects were deemed non-significant ($p>0.05$). Investigating the *Time* factor (Supplementary Material Figure S7B), similarly to the transdermal administration, a circadian rhythm was found as the major source of variability (53.5% of the *Time* factor), with marked modulations of DHEA, corticosterone, cortisol, and 11-deoxycortisol. Moreover, the effect of oral T administration was observed on the second component related to the *Time* factor (18.6%, Figure 5B). Altered levels of T, DHT and androstenedione were associated with an increase of the temporal trajectory after 2 and 4 h and a return to the basal situation.

345 Taking all this information, these results are perfectly in line with our previous study based on
346 multiway modeling [22], highlighting the marked impact of inter-individual differences and
347 intra-day variations. The decomposition of the different sources of variability can constitute an
348 objective basis for the selection of biomarkers less affected by these factors. The data mining
349 strategy was able to efficiently decompose the different sources of variation and investigate the
350 impact of T administration in the presence of confounders or other experimental factors, a
351 situation commonly found in the anti-doping routine analysis.

352

4. Conclusions

Measurement of endogenous steroid hormones in serum has been recently proven to be an interesting complementary strategy to the current urinary steroid profile for the detection of steroid-related doping abuse. In this research study, the potential of HRMS analysis in this context was evaluated by developing an UHPLC-HRMS method for the quantification of some of the most important endogenous steroids in serum related to testosterone metabolism. Its analytical performance was then compared to that of an already validated UHPLC-MS/MS method by analyzing serum samples collected during a testosterone clinical study. The results showed that for the 8 analytes that were detected with all three employed acquisition methods, concentration values measured by HRMS strongly correlated with the ones measured by MS/MS. Only in the case of low concentration steroids, such as DHT and 11-deoxycortisol, was a higher dispersion of samples in Bland-Altman plots observed in the low concentration region. This was probably due to a sensitivity gap between HRMS and MS/MS, which could be partially counterbalanced by the multiplexed t-SIM approach. The evaluation of HRMS performance was also investigated with a complementary approach based on ANOVA decomposition and multivariate analysis. Among the factors of variability, the AMOPLS showed that the influence of quantification *Method* was negligible (<2% of the total variance) in both transdermal and oral administration datasets, hence giving a biological description of testosterone intake fully in accordance with already published research based on UHPLC-MS/MS quantification.

This study noted the suitability of UHPLC-HRMS systems for the quantification of endogenous steroid hormones in serum (Quan), confirming this matrix as a promising aid to improve steroid abuse detection in the anti-doping field. **In addition, the possibility to acquire data in HRMS**

FullScan mode opens the way to untargeted studies (Qual) for new metabolites identification. Qual-Quan acquisition will be of great usefulness, in particular in the anti-doping context, where most of WADA accredited laboratories are increasing their investments on HRMS instruments.

Acknowledgments

This study was supported by the FIFA Medical Assessment and Research Centre (F-MARC). Julien Boccard and Serge Rudaz thank the Swiss National Science Foundation for its financial support (project N° 31003A_166658).

Conflict of interest

The authors declare that they have no conflict of interest

References

[1] Endogenous Anabolic Androgenic Steroids, Measurement and Reporting, Technical Document TD2016EAAS (2016) World Anti-doping Agency (WADA), Montreal. http://www.wada-ama.org. Accessed September 2016.

[2] T. Kuuranne, M. Saugy, N. Baume, Confounding factors and genetic polymorphism in the evaluation of individual steroid profiling, Br. J. Sports. Med. 48 (2014) 848-855.

[3] J. Jakobsson-Schulze, J. Lundmark, M. Garle, I. Skilving, L. Ekstrom, A. Rane, Doping Test Results Dependent on Genotype of Uridine Diphospho-Glucuronosyl Transferase 2B17, the Major Enzyme for Testosterone Glucuronidation, J. Clin. Endocr. Metab. 93 (2008) 2500-2506.

[4] A. Rane, L. Ekstrom, Androgens and doping tests: genetic variation and pit-falls, Brit. J. Clin. Pharmaco. 74 (2012) 3-15.

[5] E. Strahm, J.E. Mullen, N. Garevik, M. Ericsson, J.J. Schulze, A. Rane, L. Ekstrom, Dose-dependent testosterone sensitivity of the steroidal passport and GC-C-IRMS analysis in relation to the UGT2B17 deletion polymorphism, Drug Test. Analysis 7 (2015) 1063-1070.

[6] P. Martin-Escudero, J. Munoz-Guerra, N. Del Prado, M. Galindo Canales, M. Fuentes Ferrer, S. Vargas, A.B. Soldevilla, E. Serrano-Garde, F. Miguel-Tobal, M. Maestro de las Casas, C. Fernandez-Perez, Impact of the UGT2B17 gene deletion on the steroid profile of an athlete, Physiol. Rep. 3 (2015) doi: 10.14814/phy2.12645.

[7] F. Badoud, J. Boccard, C. Schweizer, F. Pralong, M. Saugy, N. Baume, Profiling of steroid metabolites after transdermal and oral administration of testosterone by ultra-high pressure

liquid chromatography coupled to quadrupole time-of-flight mass spectrometry, J. Steroid. Biochem. 138 (2013) 222-235.

[8] P. Van Renterghem, P. Van Eenoo, W. Van Thuyne, H. Geyer, W. Schanzer, F.T. Delbeke, Validation of an extended method for the detection of the misuse of endogenous steroids in sports, including new hydroxylated metabolites, J. Chrom. B 876 (2008) 225-235.

[9] P. Van Renterghem, P. Van Eenoo, F.T. Delbeke, Population based evaluation of a multi-parametric steroid profiling on administered endogenous steroids in single low dose, Steroids 75 (2010), 1047-1057.

[10] P. Van Renterghem, P. Van Eenoo, P.E. Sottas, M. Saugy, F.T. Delbeke, Subject-based steroid profiling and the determination of novel biomarkers for DHT and DHEA misuse in sports, Drug Test. Analysis 2 (2010) 582-588.

[11] P. Van Renterghem, P. Van Eenoo, P.E. Sottas, M. Saugy, F.T. Delbeke, A pilot study on subject-based comprehensive steroid profiling: novel biomarkers to detect testosterone misuse in sports, Clin. Endocrinol. 75 (2011) 134-140.

[12] A. Fabregat, O.J. Pozo, J. Marcos, J. Segura, R. Ventura, Use of LC-MS/MS for the open detection of steroids metabolites conjugated with glucuronic acid, Anal. Chem. 85 (2013) 5005-5014.

[13] A. Kotronoulas, A. Fabregat, I. Alfonso, T. Parella, J. Segura, R. Ventura, J. Joglar, O.J. Pozo, Synthesis andcharacterization of 6β-hydroxyandrosterone and 6β-hydroxyetiocholanolone conjugated with glucuronic acid, Drug Test. Analysis 7 (2015) 247-252.

[14] A. Kotronoulas, J. Marcos, J. Segura, R. Ventura, J. Joglar, O.J. Pozo, Ultra high performance liquid chromatography tandem mass spectrometric detection of glucluronides resistant to enzymatic hydrolysis: implications to doping control analysis, Anal. Chim. Acta 895 (2015) 35-44.

[15] J. Boccard, F. Badoud, E. Grata, S. Ouertani, M. Hanafi, G. Mazerolels, P.Lantéri, J.L. Veuthey, M. Saugy, A steroidomic approach for biomarkers discovery in doping control, Forensic Sci. Int. 213 (2011) 85-94.

[16] J. Boccard, F. Badoud, N. Jan, R. Nicoli, C. Schweizer, F. Pralong, J.L. Veuthey, N. Baume, S. Rudaz, M. Saugy, Untargeted profiling of urinary steroid metabolites after testosterone ingestion: opening new perspectives for antidoping testing, Bioanalysis 6 (2014) 2523-2536.

[17] M. Raro, M. Ibanez, R. Gil, A. Fabregat, E. Tudela, K. Deventer, R. Ventura, J. Segura, J. Marcos, A. Kotronoulas, J. Joglar, M. Farré, S. Yang, Y. Xing, P. Van Eenoo, E. Pitarch, F. Hernandez, J.V. Sancho, O.J. Pozo, Untargeted metabolomics in doping control: detection of new markers of testosterone misuse by ultrahigh performance liquid chromatography coupled to high-resolution mass spectrometry, Anal. Chem. 87 (2015) 8373-8380.

[18] F. Badoud, E. Grata, L. Perrenoud, L. Avois, M. Saugy, S. Rudaz, J.L. Veuthey, Fast analysis of doping agents in urine by ultra-high-pressure liquid chromatography-quadrupole time-of-flight mass spectrometry I. Screening analysis, J. Chrom. A 1216 (2009) 4423-4433.

[19] A. Musenga, D.A. Cowan, Use of ultra-high pressure liquid chromatography coupled to high resolution mass spectrometry for fast screening in high throughput doping control, J. Chrom. A 1288 (2013) 82-95.

447 [20] E.N. Ho, W.H. Kwok, A.S. Wong, T.S. Wan, High resolution accurate mass screening of
448 prohibited substances in equine plasma using liquid chromatography--Orbitrap mass
449 spectrometry, Drug Test. Analysis 5 (2013) 509-528.

450 [21] J.C. Dominguez-Romero, J.F. Reyes, F.J Lara-Ortega, A. Molina-Diaz, Screening and
451 confirmation capabilities of liquid chromatography-time-of-flight mass spectrometry for the
452 determination of 200 multiclass sport drugs in urine, Talanta 134 (2015) 74-88.

453 [22] F. Ponzetto, F. Mehl, J. Boccard, N. Baume, S. Rudaz, M. Saugy, R. Nicoli, Longitudinal
454 monitoring of endogenous steroids in human serum by UHPLC-MS/MS as a tool to detect
455 testosterone abuse in sports, Anal. Bioanal. Chem. 408 (2016) 705-719.

456 [23] H. Passing, L. Bablok, A new biometrical procedure for testing the equality of
457 measurements from two different analytical methods. Application of linear regression
458 procedures for method comparison studies in clinical chemistry, Part I, J. Clin. Chem. Cli.
459 Biochem. 21 (1983) 709-720.

460 [24] D.G. Altman, J.M. Bland, Measurements in medicine: the analysis of method comparison
461 studies, The Statistician 32 (1983) 307-317.

462 [25] J.M. Bland, D.G. Altman, Statistical methods for assessing agreement between two
463 methods of clinical measurement, Lancet. 1 (1986) 307-310.

464 [26] D. Giavarina, Understanding Bland Altman analysis, Biochem. Med. (Zagreb) 25 (2015) 141-
465 151.

[27] S. Rudaz, S. Souverain, C. Schelling, M. Deelers, A. Klomp, A. Norris, B. Ariano, J.-L. Veuthey, Development and validation of a heart-cutting LC-MS method for the determination of process related substances in cetirizine material, Anal. Chim. Acta 492 (2003) 271-282.

[28] Food and Drug Administration, Guidance for Industry. Bioanalytical Method Validation, FDA, Center for Drug Evaluation and Research. Rockeville, 2001.

[29] J. Boccard, S. Rudaz, Exploring Omics data from designed experiments using Analysis of Variance Multiblock Orthogonal Partial Least Squares, Anal. Chim. Acta 920 (2016) 18-28.

Tables

Table 1 - Description of target analytes and detection parameters

Analyte	Elemental Composition	Retention Time [min]	Observed Ion Mass (monoisotopic) [m/z]	Theoretical Ion Mass (monoisotopic) [m/z]	Mass Error [ppm]	MSX Count*	Time Window [min]
Testosterone	$C_{19}H_{28}O_2$	4.61	289.12579	289.21621	0.93	2	4.1 - 4.7
Epitestosterone	$C_{19}H_{28}O_2$	5.31	289.21582	289.21621	1.35	8	4.6 - 5.5
Androstenedione	$C_{19}H_{26}O_2$	5.13	287.20026	287.20056	0.84	8	4.6 - 5.5
Progesterone	$C_{21}H_{30}O_2$	6.85	315.23157	315.23186	0.54	2	6.2 - 7.2
17α-Hydroxyprogesterone	$C_{21}H_{30}O_3$	5.26	331.22641	331.22677	1.27	8	4.6 - 5.5
DHEA	$C_{19}H_{28}O_2$	5.10	271.20535°	271.20564°	0.96	8	4.6 - 5.5
DHT	$C_{19}H_{30}O_2$	5.60	291.23175	291.23186	1.10	2	5.3 - 5.7
Corticosterone	$C_{21}H_{30}O_4$	3.63	347.22122	347.22169	0.72	3	3.1 - 4.1
Cortisol	$C_{21}H_{30}O_5$	2.77	363.21619	363.21660	1.57	2	2.3 - 3.0
Deoxycorticosterone	$C_{21}H_{30}O_3$	4.86	331.22638	331.22677	1.09	2	4.5 - 5.0
11-Deoxycortisol	$C_{21}H_{30}O_4$	3.77	347.22124	347.22169	1.01	3	3.1 - 4.1

* Also taking into account the deuterated internal standard for each target analyte

° $[M\text{-}H_2O]^+$

Table 2 – Estimated lower limit of quantification (LLOQ) for target analytes with FS and t-SIM acquisition modes in comparison with the MS/MS method

Compound	LLOQ [pg/mL]		
	FS	t-SIM	MS/MS[22]
Testosterone	100	50	20
Epitestosterone	100	50	20
Androstenedione	100	50	50
17α-Hydroxyprogesterone	500	250	100
Progesterone	100	50	15
DHEA	500	500	500
DHT	500	250	50
Corticosterone	100	100	100
Cortisol	1000	1000	1000
Deoxycorticosterone	500	250	25
11-Deoxycortisol	250	250	25

Figures captions

Figure 1 Chromatograms of 17α-hydroxyprogesterone obtained with three different HRMS acquisition modes: (A) Full scan, (B) t-SIM without multiplexing and (C) t-SIM with MSX count of 8

Figure 2 Passing-Bablok regressions and Bland-Altman plots for the comparison of testosterone quantification results. (A, B) t-SIM vs SRM and (C, D) FS vs SRM

Figure 3 Passing-Bablok regressions and Bland-Altman plots for the comparison of DHT quantification results. (A, B) t-SIM vs SRM and (C, D) FS vs SRM

Figure 4 AMOPLS modeling for transdermal administration data (30.3% of explained variance): (A) relative contributions of *volunteer*, *time* and *method* effects and their interactions; (B) score and loadings plots of predictive component related to the *time* factor. C: control, P: patch.

Figure 5 AMOPLS modeling for oral administration data (18.6% of explained variance): (A) relative contributions of *volunteer*, *time* and *method* effects and their interactions; (B) score and loadings plots of predictive component related to the *time* factor. C: control, O: oral.

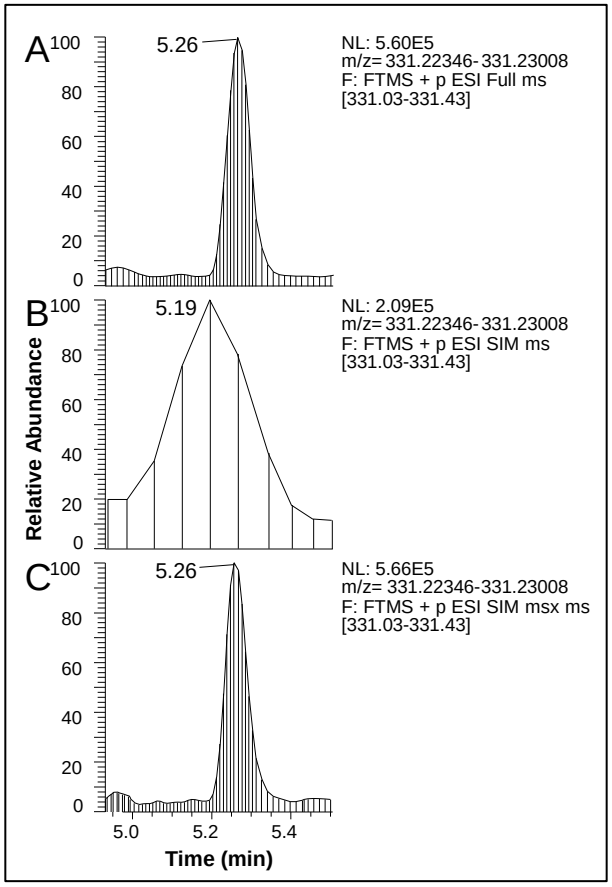
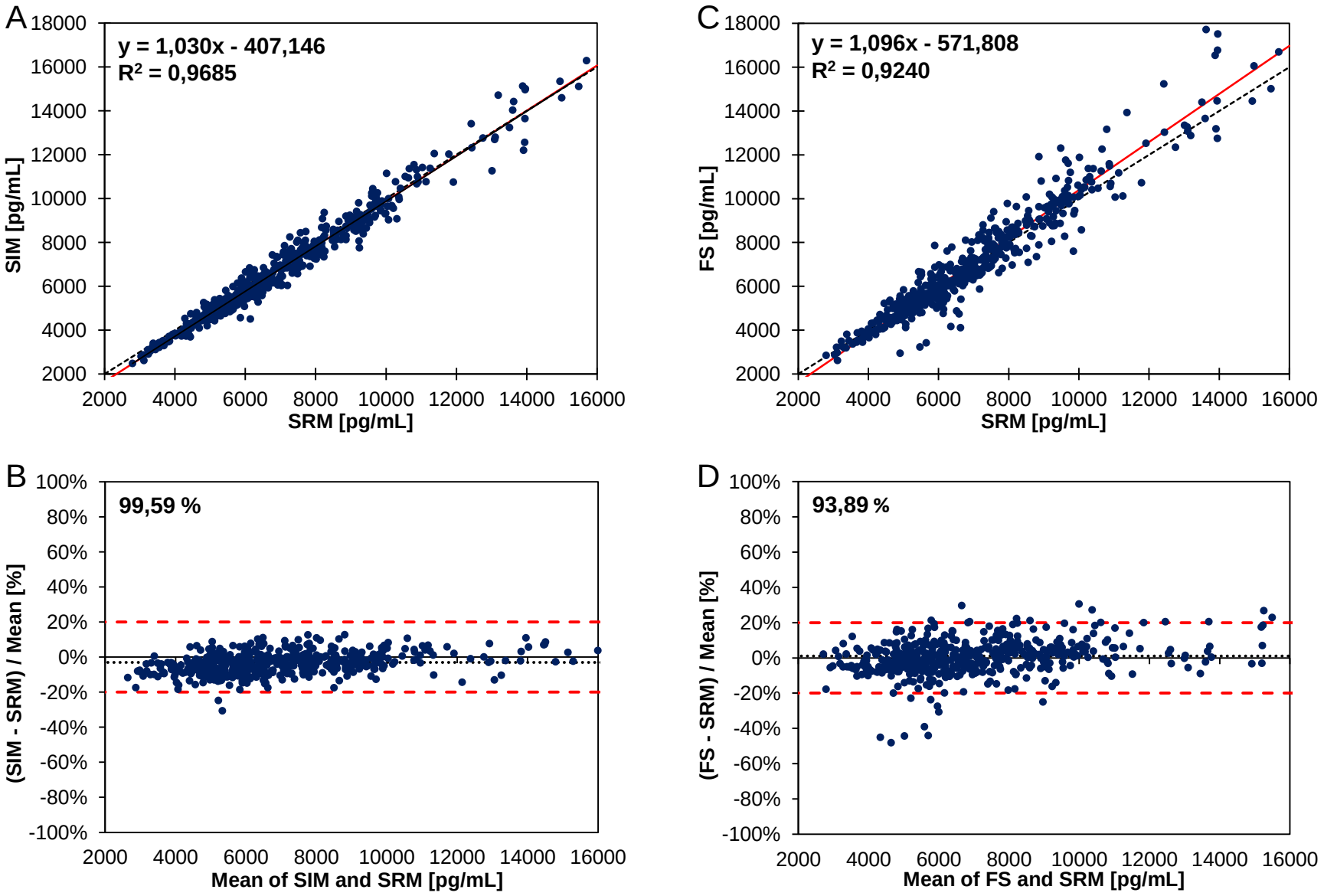
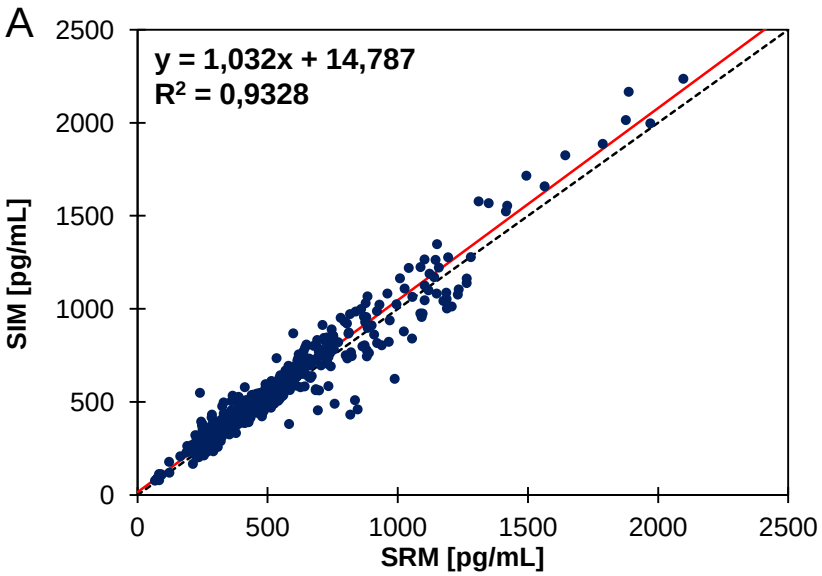
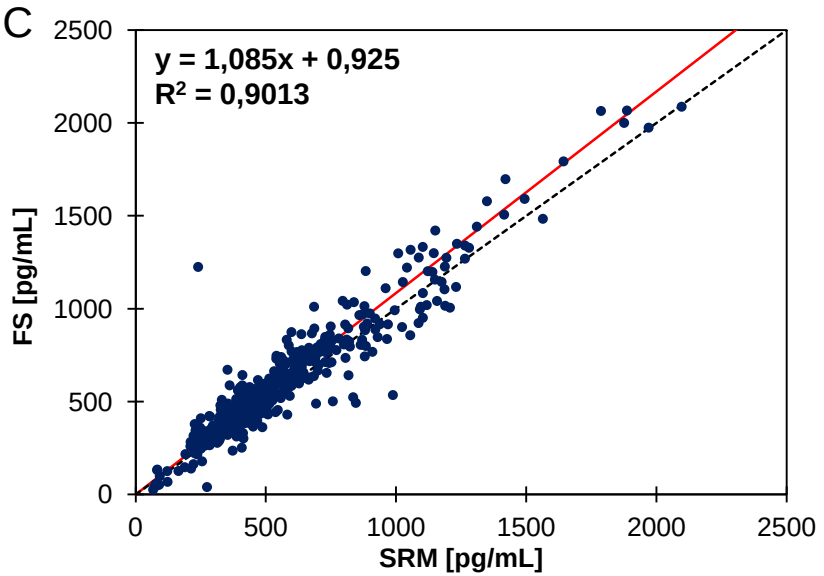
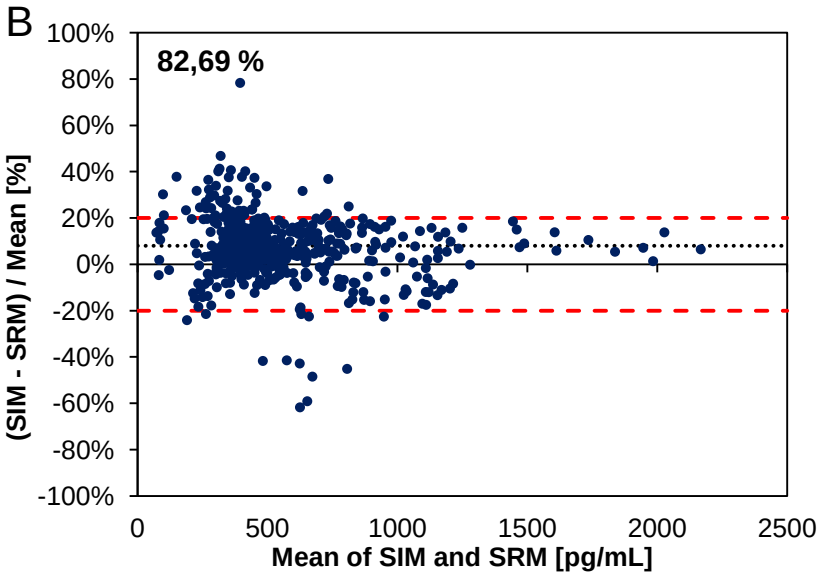

A
B
C
5.26
5.19
5.26
NL: 5.60E5
m/z= 331.22346-331.23008
F: FTMS + p ESI Full ms
[331.03-331.43]
NL: 2.09E5
m/z= 331.22346-331.23008
F: FTMS + p ESI SIM ms
[331.03-331.43]
NL: 5.66E5
m/z= 331.22346-331.23008
F: FTMS + p ESI SIM msx ms
[331.03-331.43]
Relative Abundance
100
80
60
40
20
0
5.0
5.2
5.4
Time (min)

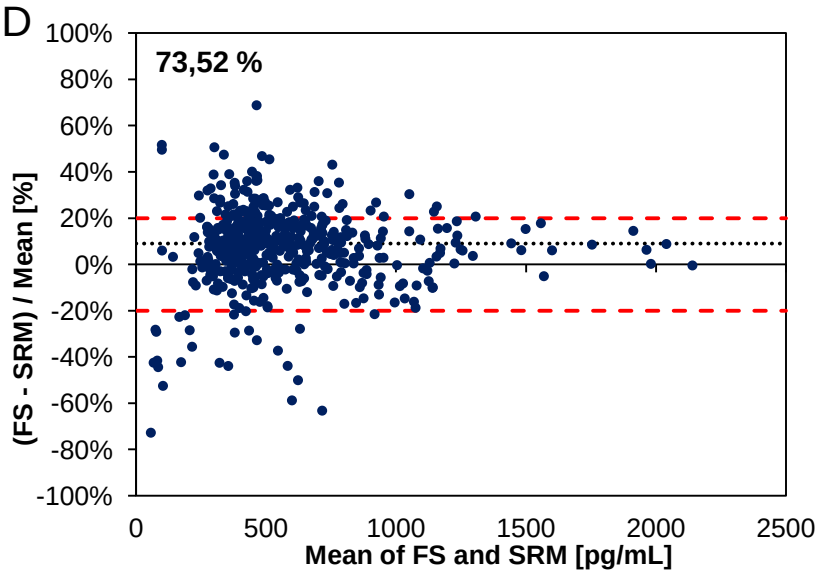
Figure 1


A
y = 1,030x - 407,146
R² = 0,9685
SIM [pg/mL]
SRM [pg/mL]
C
y = 1,096x - 571,808
R² = 0,9240
FS [pg/mL]
SRM [pg/mL]
B
99,59 %
(SIM - SRM) / Mean [%]
Mean of SIM and SRM [pg/mL]
D
93,89 %
(FS - SRM) / Mean [%]
Mean of FS and SRM [pg/mL]

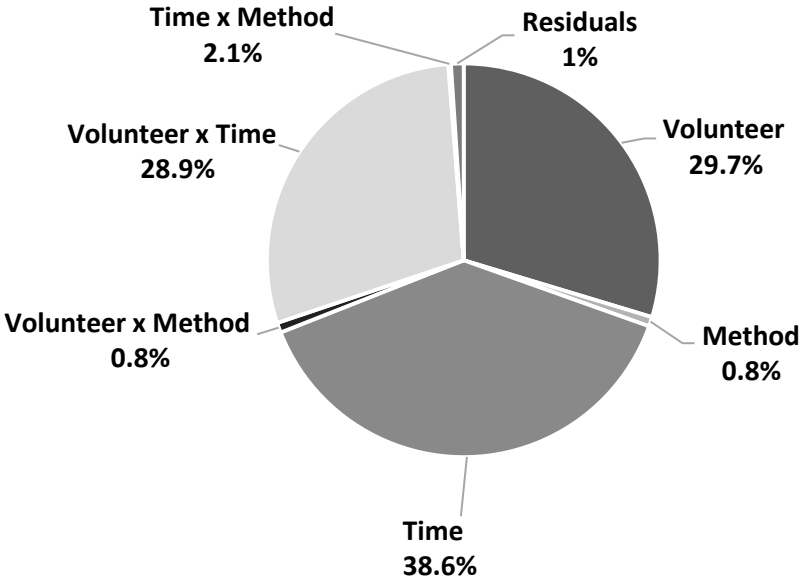

A
y = 1,032x + 14,787
R² = 0,9328
SIM [pg/mL]
SRM [pg/mL]
0
500
1000
1500
2000
2500


C
y = 1,085x + 0,925
R² = 0,9013
FS [pg/mL]
SRM [pg/mL]
0
500
1000
1500
2000
2500

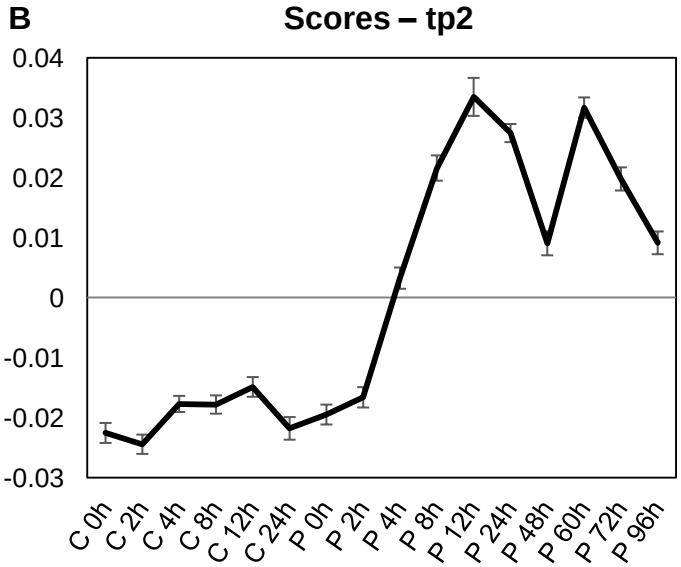

B
82,69 %
(SIM - SRM) / Mean [%]
Mean of SIM and SRM [pg/mL]
100%
80%
60%
40%
20%
0%
-20%
-40%
-60%
-80%
-100%
0
500
1000
1500
2000
2500

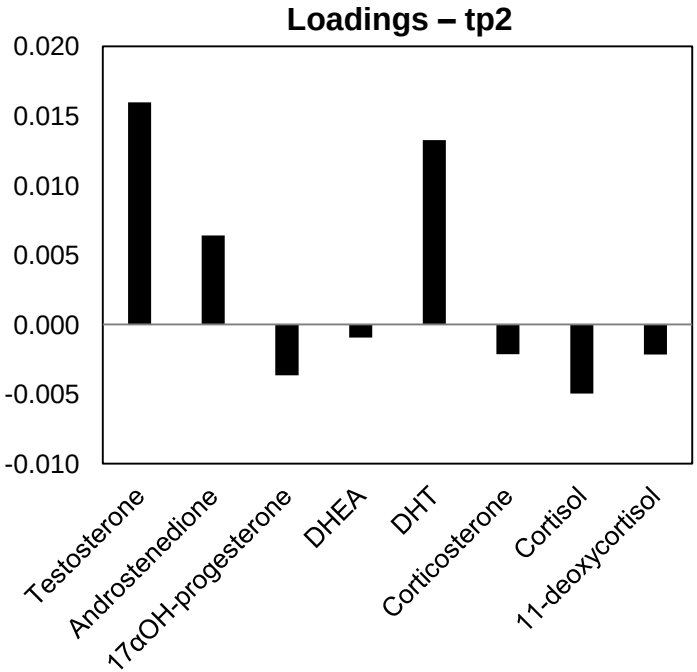

D
73,52 %
(FS - SRM) / Mean [%]
Mean of FS and SRM [pg/mL]
100%
80%
60%
40%
20%
0%
-20%
-40%
-60%
-80%
-100%
0
500
1000
1500
2000
2500

A

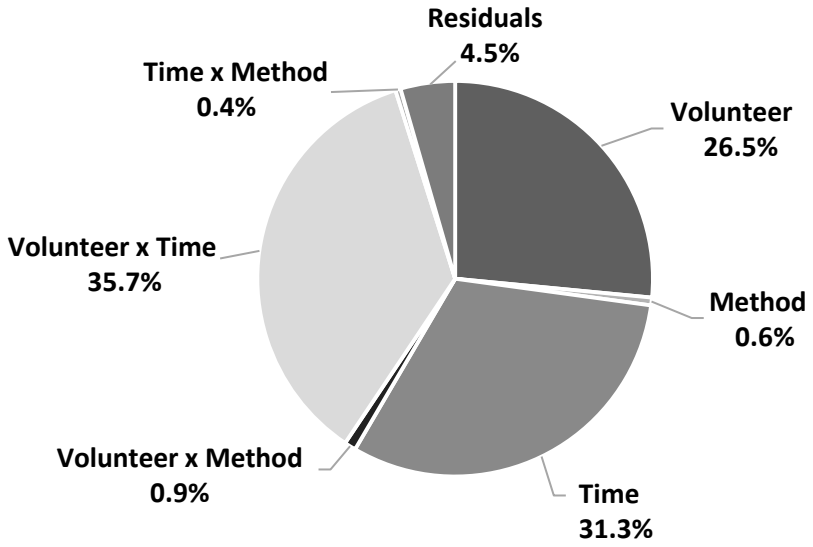

Time x Method
2.1%
Residuals
1%
Volunteer
29.7%
Volunteer x Time
28.9%
Volunteer x Method
0.8%
Method
0.8%
Time
38.6%

B

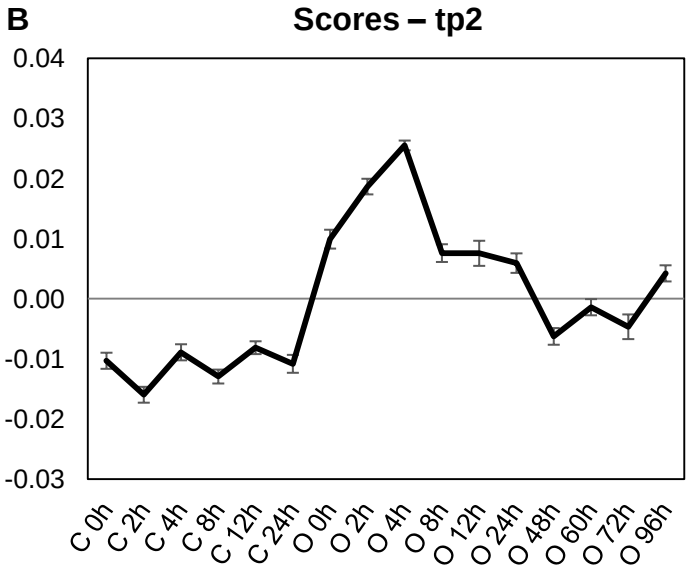

Scores – tp2
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
C 0h
C 2h
C 4h
C 8h
C 12h
C 24h
P 0h
P 2h
P 4h
P 8h
P 12h
P 24h
P 48h
P 60h
P 72h
P 96h


Loadings – tp2
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
Testosterone
Androstenedione
17αOH-progesterone
DHEA
DHT
Corticosterone
Cortisol
11-deoxycortisol

A


Residuals
4.5%
Time x Method
0.4%
Volunteer
26.5%
Volunteer x Time
35.7%
Method
0.6%
Volunteer x Method
0.9%
Time
31.3%

B


Scores – tp2
0.04
0.03
0.02
0.01
0.00
-0.01
-0.02
-0.03
C 0h
C 2h
C 4h
C 8h
C 12h
C 24h
O 0h
O 2h
O 4h
O 8h
O 12h
O 24h
O 48h
O 60h
O 72h
O 96h

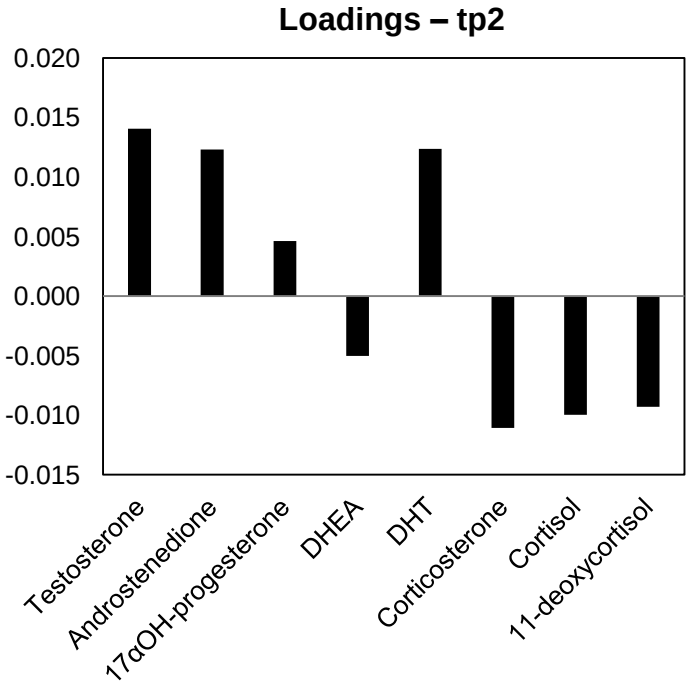

Loadings – tp2
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
Testosterone
Androstenedione
17αOH-progesterone
DHEA
DHT
Corticosterone
Cortisol
11-deoxycortisol

Table S1 Internal Standard Mix composition (final conc. in serum)

Compound	Concentration (ng/mL)
Testosterone-d3	0,2
Epitestosterone-d3	0,5
Androstenedione-d7	5
Progesterone-d9	2
17α-Hydroxyprogesterone-d8	10
DHEA-d5	4

Compound	Concentration (ng/mL)
DHT-d3	2
Corticosterone-d8	5
Cortisol-d4	2,5
Deoxyorticosterone-d8	2
11-Deoxycortisol-d2	1

Table S2 Calibration samples composition (final conc. in serum)

Compound	Concentration (pg/mL)							
	Cal1	**Cal2**	**Cal3**	**Cal4**	**Cal5**	**Cal6**	**Cal7**	**Cal8**
Testosterone	20	50	100	500	1000	5000	10000	25000
Epitestosterone	20	50	100	500	1000	2500	10000	
Androstenedione	50	100	500	1000	2500	5000	10000	25000
Progesterone	15	25	50	100	500	2500	10000	25000
17α-Hydroxyprogesterone	100	250	500	1000	2500	500	10000	25000
DHEA	500	2500	5000	10000	25000			
DHT	50	250	500	750	1000	2500	5000	10000
Corticosterone		100	500	1000	5000	25000	50000	100000
Cortisol	1000	2500	10000	25000	100000	200000	300000	400000
Deoxycorticosterone		25	250	500	1000	2500	10000	
11-Deoxycortisol		25	250	500	1000	2500	5000	

Table S3 Description of labeled internal standards and detection parameters

Analyte	Elemental Composition	Retention Time [min]	Observed Ion Mass (monoisotopic) [m/z]	Theoretical Ion Mass (monoisotopic) [m/z]	Mass Error [ppm]	MSX Count	Time Window [min]
Testosterone-d3	C19H25D3O2	4.58	292.23484	292.23431	1.81	2	4.1 - 4.7
Epitestosterone-d3	C19H25D3O2	5.28	292.23484	292.23431	1.81	8	4.6 - 5.5
Androstenedione-d7	C19H19D7O2	5.10	294.24429	294.24384	1.53	8	4.6 - 5.5
Progesterone-d9	C21H21D9O2	6.81	324.28835	324.28784	1.57	2	6.2 - 7.2
17α-Hydroxyprogesterone-d8	C21H22D8O3	5.24	339.27679	339.27628	1.50	8	4.6 - 5.5
DHEA-d5	C19H23D5O2	5.07	276.23703	276.23706	-0.11	8	4.6 - 5.5
DHT-d3	C19H27D3O2	5.58	294.25069	294.25027	1.43	2	5.3 - 5.7
Corticosterone-d8	C21H22D8O4	3.60	355.27190	355.27155	0.99	3	3.1 - 4.1
Cortisol-d4	C21H26D4O5	2.77	367.24171	367.24142	0.79	2	2.3 - 3.0
Deoxycorticosterone-d8	C21H22D8O3	4.82	339.27699	339.27634	1.92	2	4.5 - 5.0
11-Deoxycortisol-d2	C21H28D2O4	3.75	349.23424	349.23384	1.15	3	3.1 - 4.1

° $[M\text{-}H_2O]^+$

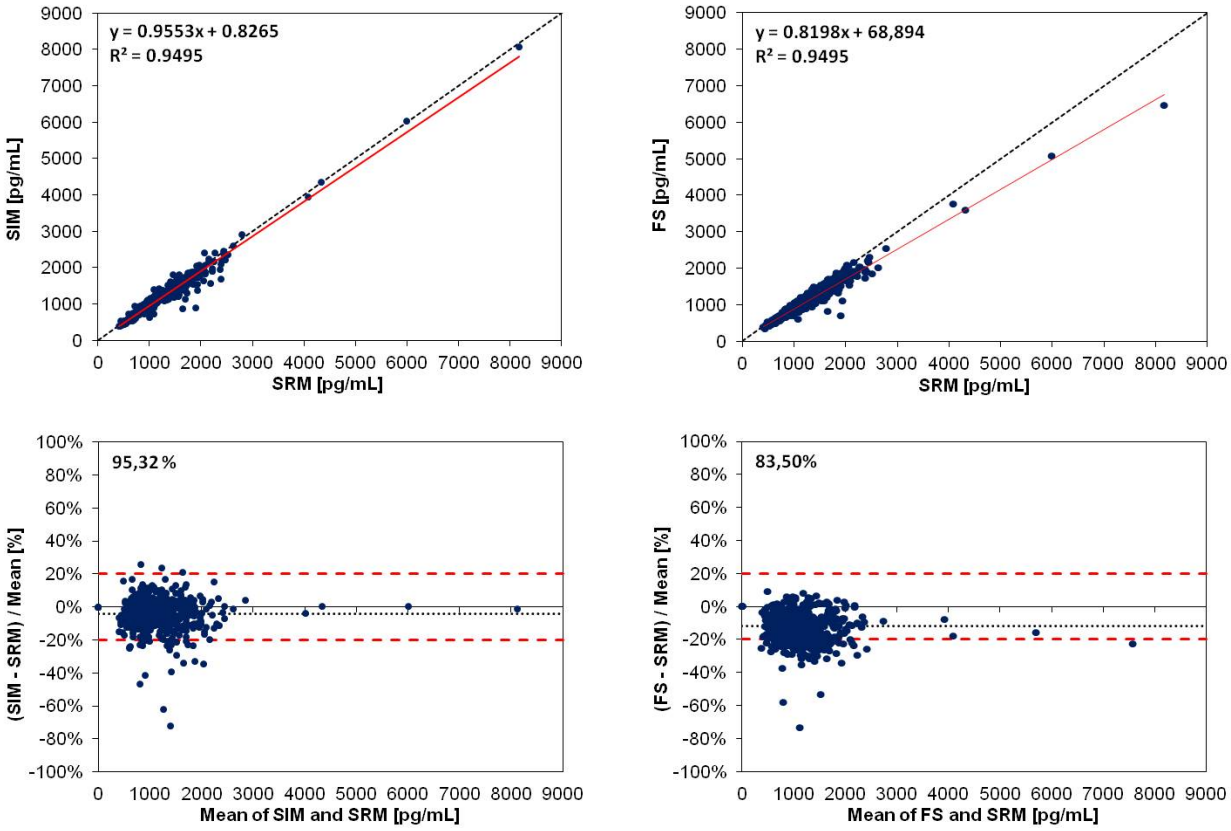


Figure S1 – Passing-Bablok regressions and Bland-Altman plots for androstenedione

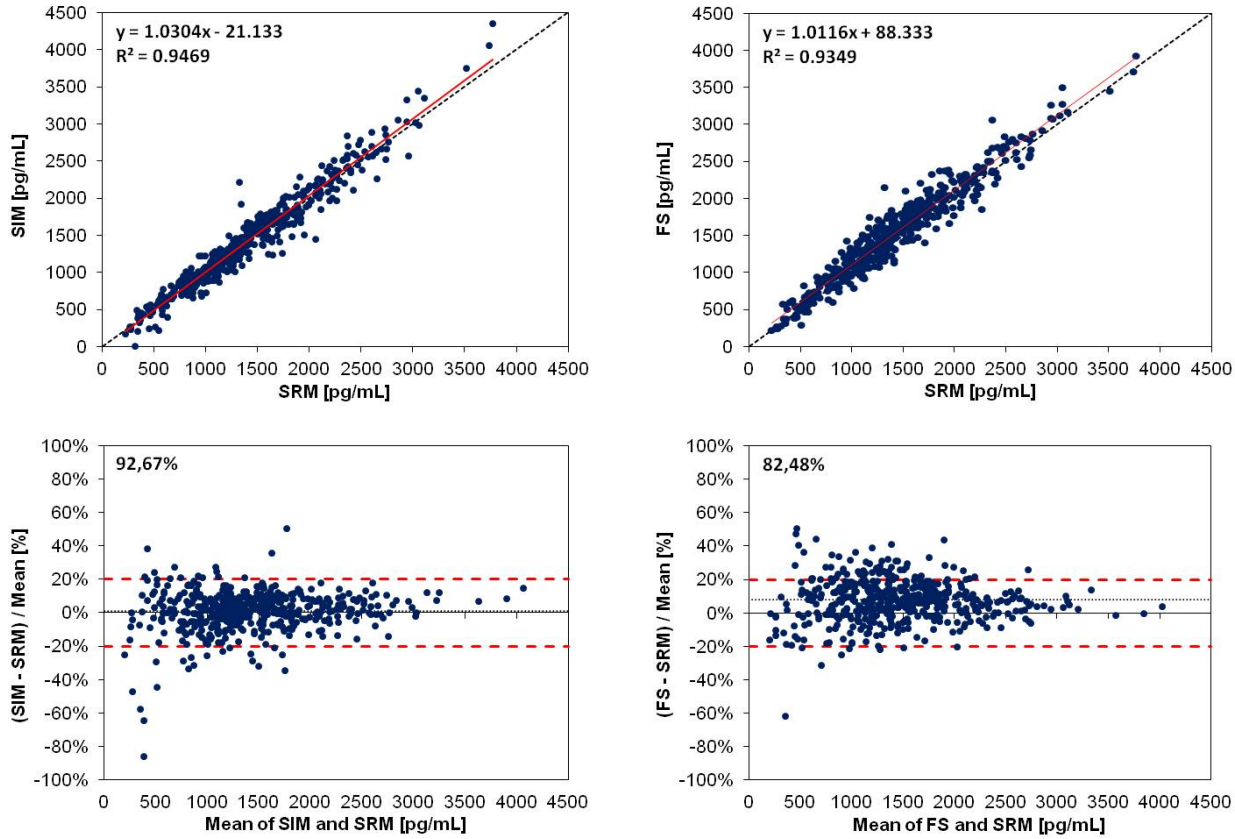


Figure S2 - Passing-Bablok regressions and Bland-Altman plots for 17α-hydroxyprogesterone

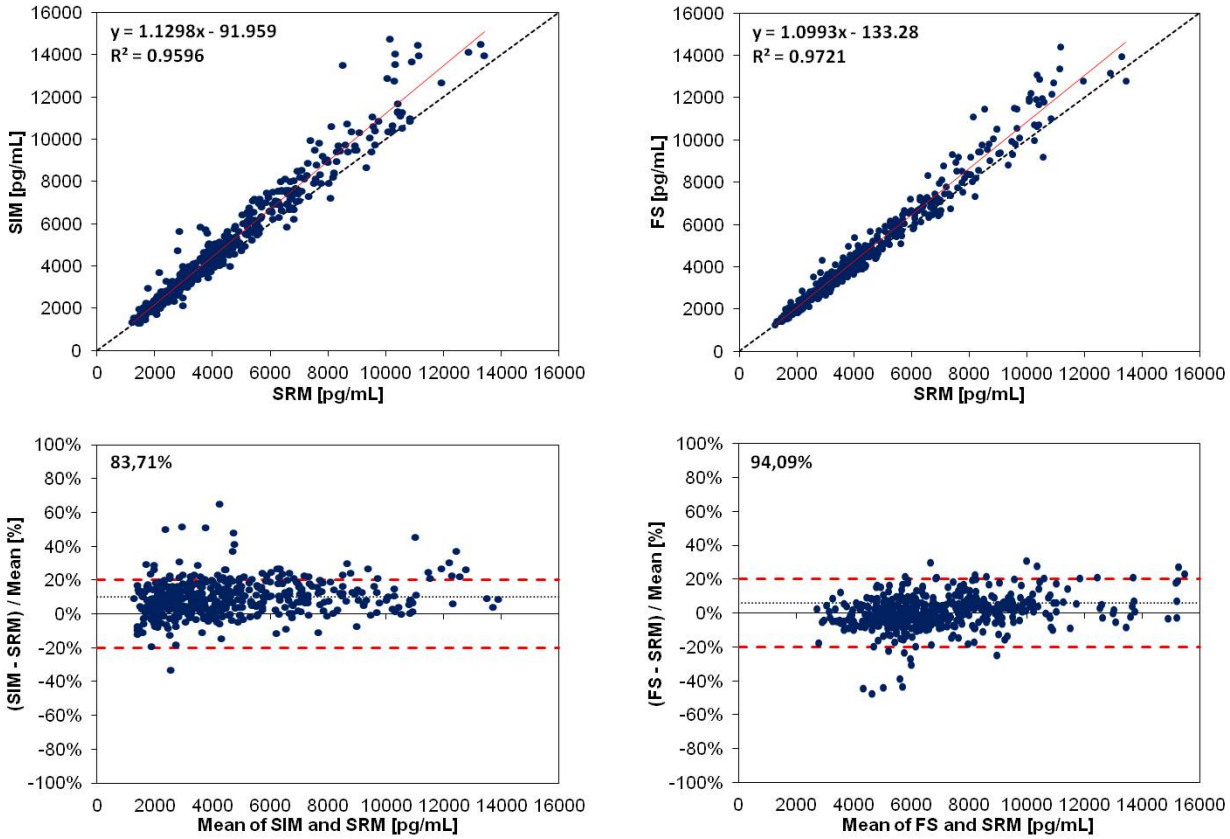


Figure S3 - Passing-Bablok regressions and Bland-Altman plots for DHEA

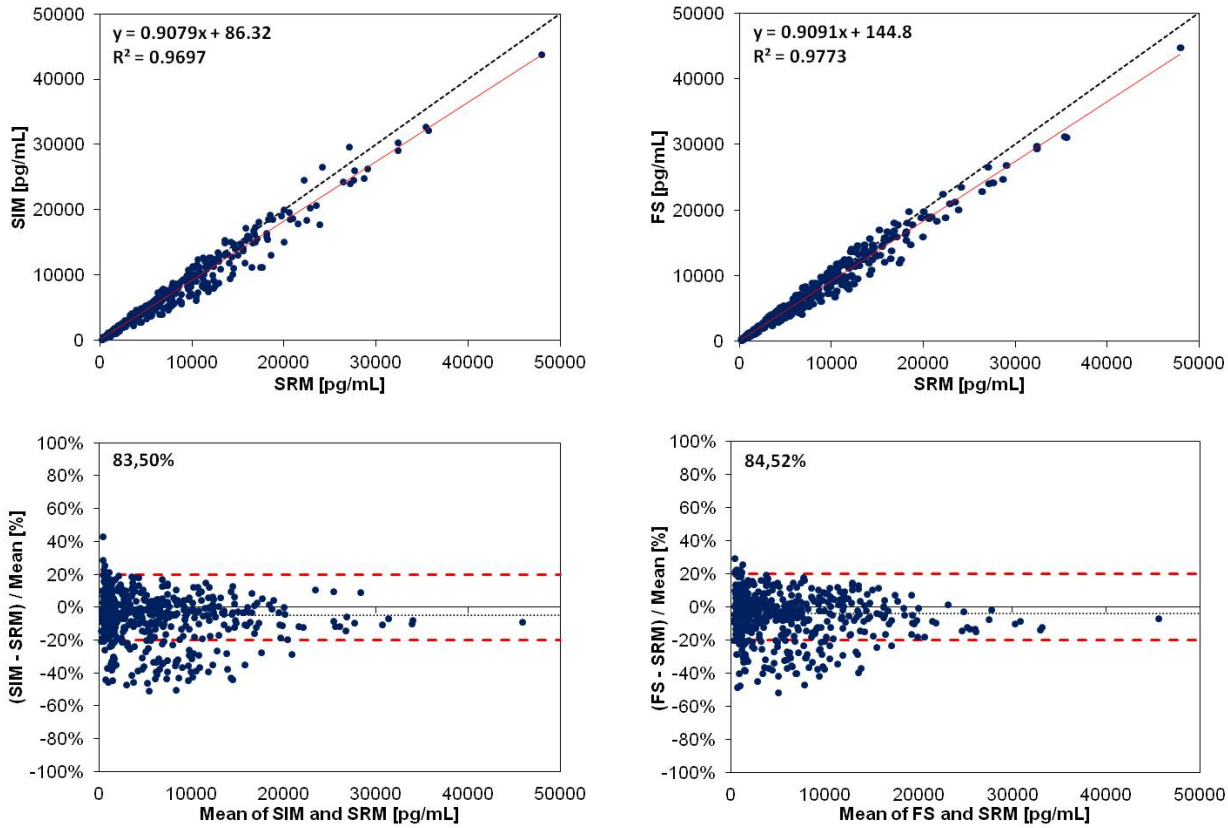


Figure S4 - Passing-Bablok regressions and Bland-Altman plots for corticosterone

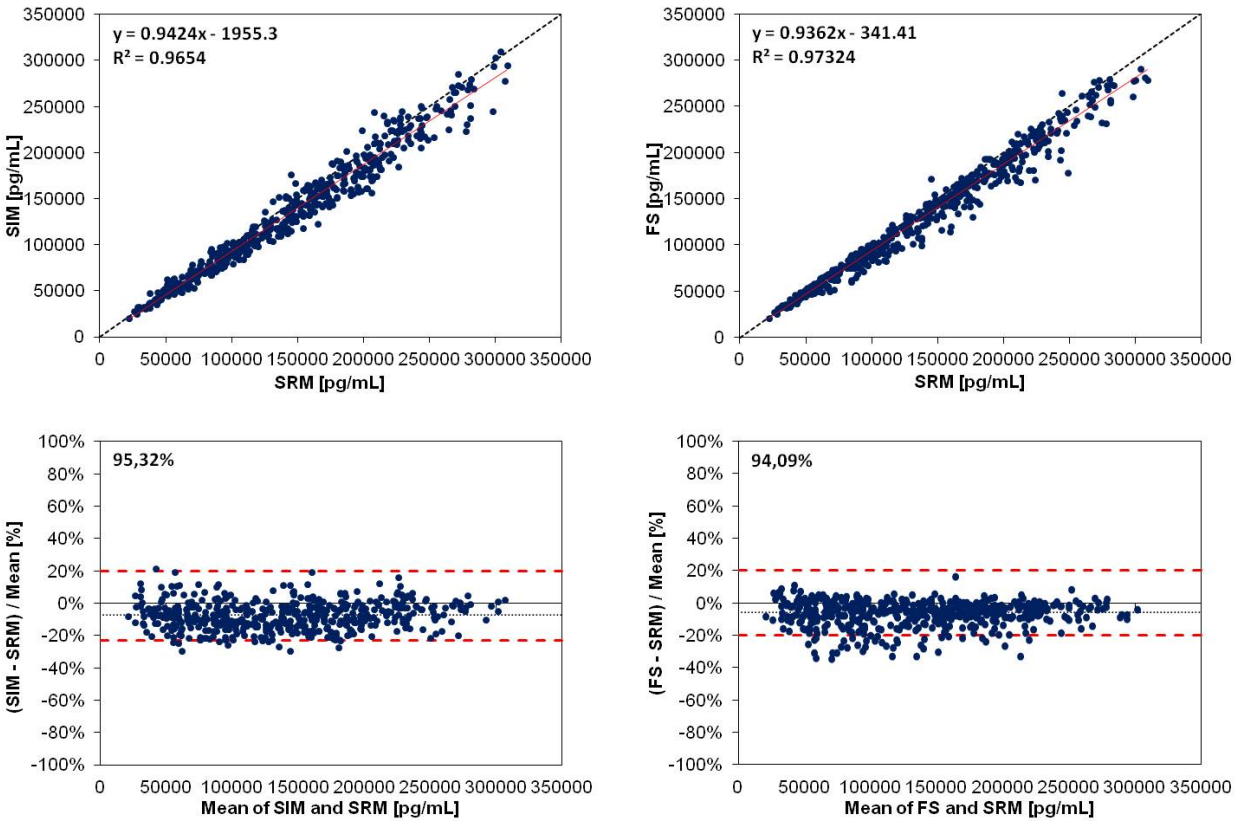


Figure S5 - Passing-Bablok regressions and Bland-Altman plots for cortisol

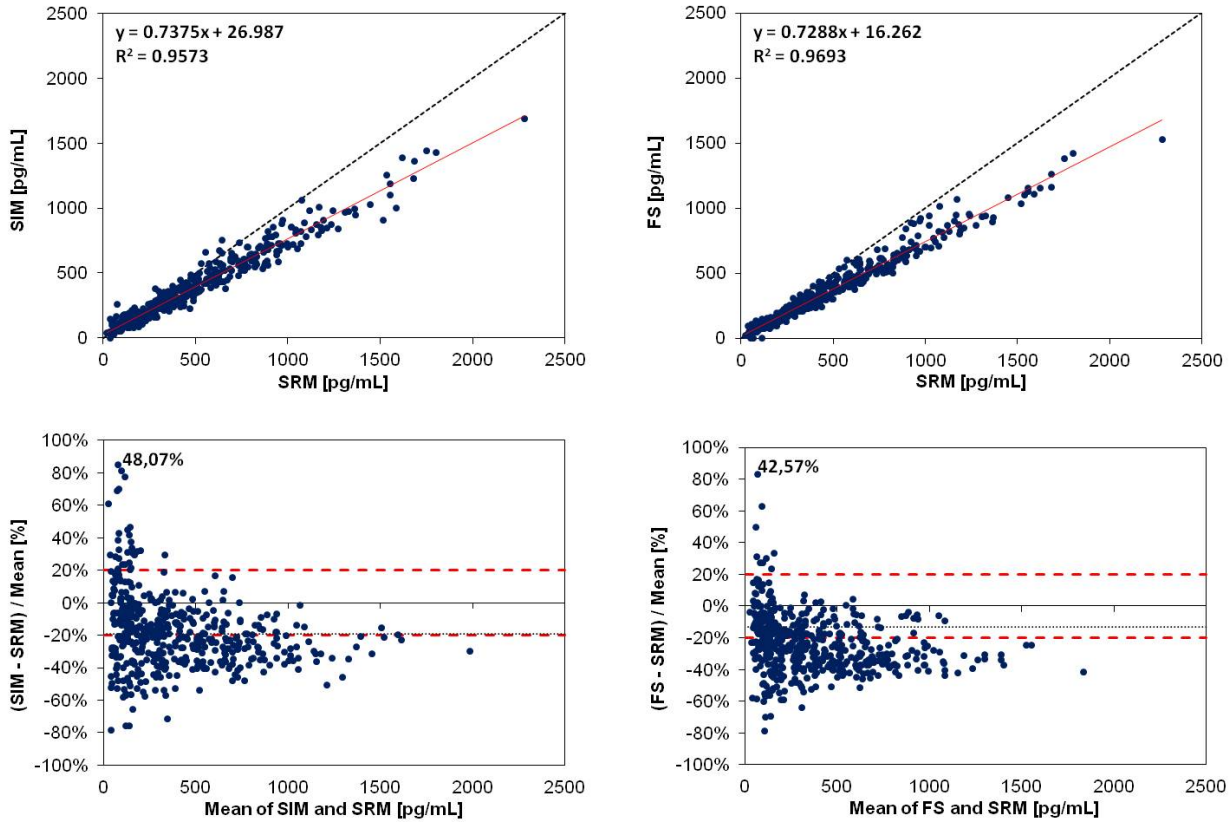


Figure S6 - Passing-Bablok regressions and Bland-Altman plots for 11-deoxycortisol

Table S4 –Trueness, repeatability and intermediate precision calculated for all analytes at each concentration level with t-SIM acquisition mode. Highlighted in grey LLOQ values, defined as the lowest concentration at which measured accuracy and precision are <20%.

Analyte	Concentration (pg/mL)	Trueness (%)*	Precision	
			Repeatability (%, n=2)	IP (%, n=12)*
Testosterone	20	127.0	5.5	12.9
	50	94.5	11.1	12.0
	100	94.6	4.3	5.6
	500	91.4	9.3	9.3
	1000	92.8	7.4	7.7
	5000	105.2	13.1	13.1
	10000	91.8	7.7	7.7
	25000	104.0	7.9	7.9
Epitestosterone	20	-	-	-
	50	116.4	14.0	14.0
	100	111.6	11.1	11.1
	500	89.4	5.0	5.1
	1000	84.6	7.0	7.1
	2500	94.9	6.4	6.8
	10000	157.8	14.0	14.0
Androstenedione	50	100.3	13.3	14.0
	100	95.1	7.6	12.4
	500	100.9	8.6	8.6
	1000	98.0	9.5	9.5
	2500	101.1	5.0	5.0
	5000	103.7	5.2	8.3
	10000	95.6	9.0	9.0
	25000	101.8	6.9	6.9
Progesterone	15	98.5	48.0	48.0
	25	93.5	23.6	23.6
	50	93.0	9.2	10.0
	100	99.0	16.6	16.6
	500	103.5	19.1	19.1
	2500	102.2	12.5	13.7
	10000	101.0	13.0	13.0
	25000	100.1	11.2	11.2

17α-hydroxyprogesterone	100	102.0	23.8	23.8
	250	90.1	10.3	10.3
	500	95.3	12.1	12.1
	1000	102.5	9.0	10.2
	2500	97.0	7.0	10.3
	5000	95.4	4.5	7.4
	10000	97.0	9.9	11.1
	25000	101.7	6.9	6.9
DHEA	500	117.4	6.6	6.6
	2500	87.5	6.9	6.9
	5000	95.8	8.8	9.0
	10000	89.1	7.1	7.1
	25000	106.2	5.2	5.2
DHT	50	101.3	23.4	23.4
	250	85.1	4.8	5.0
	500	108.2	6.5	8.2
	750	105.6	6.6	7.4
	1000	108.2	3.0	6.2
	2500	98.0	9.1	9.1
	5000	92.4	6.3	6.3
	10000	106.0	4.5	4.9
Corticosterone	100	104.1	16.7	17.3
	1000	100.0	10.9	11.3
	5000	96.0	11.0	11.0
	25000	100.2	11.6	11.6
	50000	105.2	7.5	9.4
	100000	95.8	8.1	10.8
Cortisol	1000	108.8	8.7	9.9
	2500	89.2	6.4	6.4
	10000	98.7	7.8	9.4
	25000	101.4	5.9	9.6
	100000	97.2	7.5	9.2
	200000	99.0	11.1	11.1
	300000	94.1	9.6	9.6
	400000	105.2	3.2	4.0

Deoxycorticosterone	25	-	-	-
	250	101.1	13.0	17.3
	500	118.7	16.5	19.3
	1000	119.4	8.3	16.4
	2500	119.0	2.8	10.4
	10000	87.5	3.9	6.0
11-deoxycortisol	25	153.9	36.1	36.1
	250	86.8	7.2	9.1
	500	86.6	6.8	12.4
	1000	83.6	8.5	8.7
	2500	90.4	16.2	16.2
	5000	87.7	7.8	11.7

*IP: Intermediate Precision

- : analyte not detected

Table S5 –Trueness, repeatability and intermediate precision calculated for all analytes at each concentration level with FS acquisition mode. Highlighted in grey LLOQ values, defined as the lowest concentration at which measured accuracy and precision are <20%.

Analyte	Concentration (pg/mL)	Trueness (%)*	Precision	
			Repeatability (%, n=2)	IP (%, n=12)*
Testosterone	20	146.2	19.2	35.8
	50	110.4	9.6	29.9
	100	92.1	16.3	16.3
	500	86.2	13.7	14.6
	1000	87.8	11.4	11.4
	5000	109.2	12.4	12.4
	10000	91.9	14.3	14.3
	25000	106.3	4.6	6.0
Epitestosterone	20	-	-	-
	50	-	-	-
	100	115.6	19.1	19.1
	500	97.2	18.6	18.6
	1000	90.2	15.8	15.8
	2500	99.0	14.5	14.5
	10000	101.0	8.1	8.1
Androstenedione	50	115.8	34.2	39.2
	100	100.5	17.8	17.8
	500	94.2	15.9	15.9
	1000	100.5	3.2	5.6
	2500	100.1	4.9	6.3
	5000	101.5	4.2	5.5
	10000	96.5	7.6	7.6
	25000	101.2	2.3	3.2
Progesterone	15	121.6	59.7	59.7
	25	101.7	31.1	43.3
	50	106.5	34.4	34.4
	100	93.7	9.2	9.7
	500	94.8	4.1	7.5
	2500	95.9	5.0	5.0
	10000	98.2	6.5	6.5
	25000	100.6	7.2	7.2

17α-hydroxyprogesterone	100	-	-	-
	250	89.4	12.1	27.0
	500	86.9	16.5	16.7
	1000	109.3	7.6	11.7
	2500	94.5	8.8	8.8
	5000	98.2	2.0	6.3
	10000	95.5	8.6	8.6
	25000	101.3	8.7	8.7
DHEA	500	117.9	2.6	2.7
	2500	87.0	8.8	8.8
	5000	93.8	3.6	6.7
	10000	87.3	3.5	4.5
	25000	107.3%	4.4	4.4
DHT	50	109.6	21.2	21.2
	250	79.7	4.4	5.0
	500	105.1	10.3	10.3
	750	101.9	7.5	8.6
	1000	110.4	8.3	9.2
	2500	97.1	6.7	7.1
	5000	92.3	6.2	6.2
	10000	105.2	5.2	5.2
Corticosterone	100	101.9	13.9	15.8
	1000	100.6	15.4	15.4
	5000	94.2	9.2	10.6
	25000	100.1	12.6	12.6
	50000	105.1	8.3	9.1
	100000	96.9	6.8	10.4
Cortisol	1000	112.9	5.0	7.1
	2500	88.2	3.4	3.4
	10000	98.1	5.4	5.8
	25000	101.0	2.4	5.6
	100000	99.2	3.4	5.7
	200000	98.1	5.4	5.6
	300000	93.5	7.8	7.8
	400000	106.2	5.2	5.5

Deoxycorticosterone	25	-	-	-
	250	83.6	9.5	22.1
	500	111.2	15.7	15.7
	1000	119.1	17.7	17.9
	2500	119.8	18.7	18.7
	10000	92.9	5.6	5.6
11-deoxycortisol	25	162.5	18.5	26.3
	250	83.7	4.9	10.1
	500	83.3	6.4	9.7
	1000	81.6	8.5	9.8
	2500	88.6	20.0	20.0
	5000	89.2	11.0	11.0

*IP: Intermediate Precision

- : analyte not detected

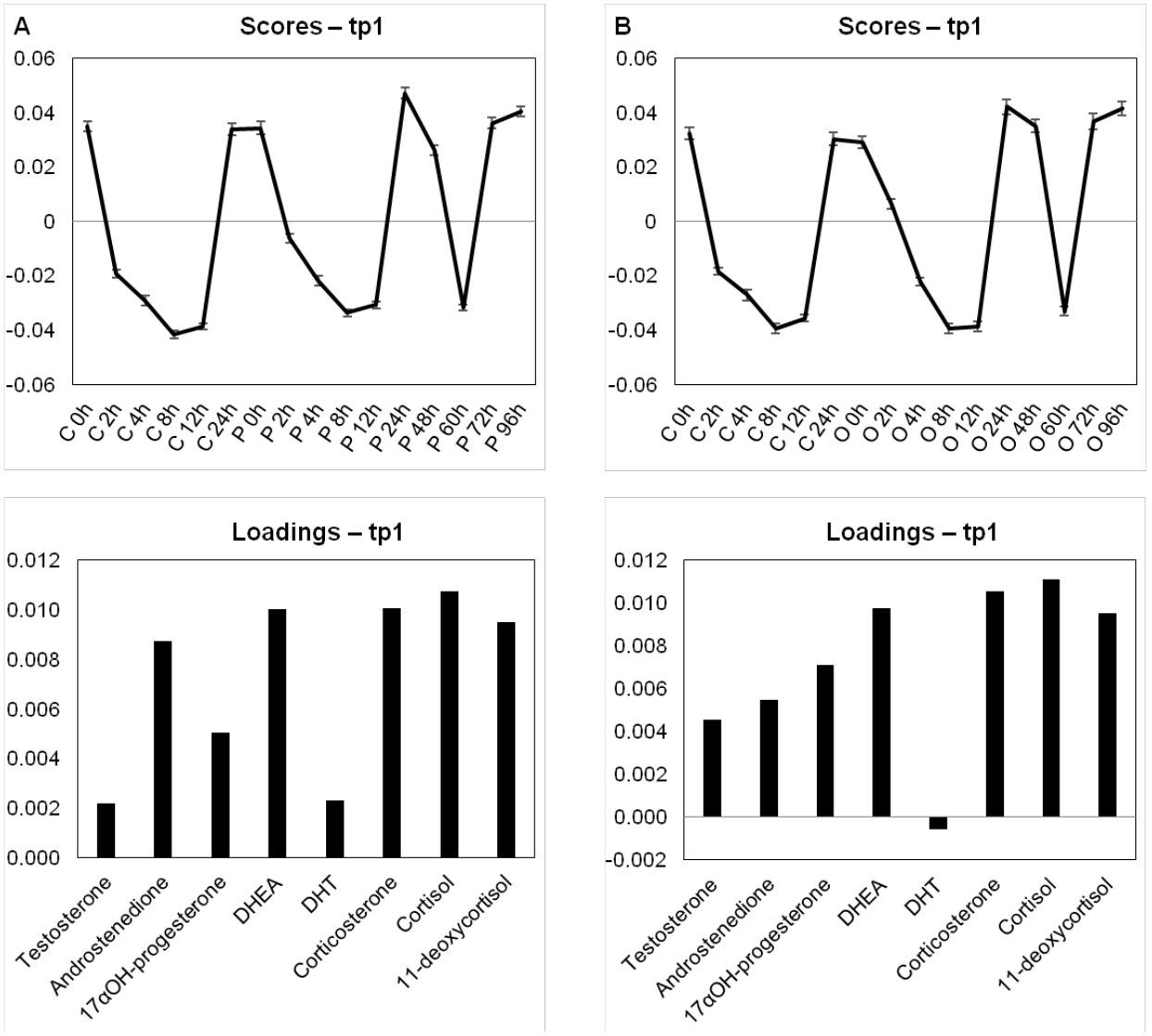


Figure S7 - Score and loadings plots of time modes of tp 1 obtained from AMOPLS modeling for (A) transdermal administration data (47.9% of explained variance) and (B) oral administration data (53.5% explained variance). C: control, P: patch, O: oral