Forensic Science International 213 (2011) 85–94

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

A steroidomic approach for biomarkers discovery in doping control☆

Julien Boccard [a,b,1], Flavia Badoud [a,b,c,1], Elia Grata [b,c], Samia Ouertani [d,e], Mohamed Hanafi [d,e], Gérard Mazerolles [f], Pierre Lantéri [g], Jean-Luc Veuthey [a,b], Martial Saugy [b,c], Serge Rudaz [a,b,*]

[a] School of Pharmaceutical Sciences, University of Geneva and Lausanne, 20 Bd d'Yvoy, 1211 Geneva 4, Switzerland
[b] Swiss Centre for Applied Human Toxicology, University of Geneva, CMU, Rue Michel-Servet 1, 1211 Geneva 4, Switzerland
[c] Swiss Laboratory for Doping Analyses, University Center of Legal Medecine, Geneva and Lausanne, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Ch. des Croisettes 22, 1066 Epalinges, Switzerland
[d] ONIRIS, Unité de Sensométrie et de Chimiométrie, Nantes, France
[e] Université Nantes Angers Le Mans, France
[f] INRA, UMR 1083, F-34060 Montpellier, France
[g] Laboratory of Analytical Sciences, University Claude Bernard Lyon I, CNRS, Villeurbanne, France

ARTICLE INFO

Article history:
Received 19 April 2011
Received in revised form 8 July 2011
Accepted 12 July 2011
Available online 9 August 2011

Keywords:
Steroidomics
UHPLC-QTOF-MS^E
Chemometrics
Doping control

ABSTRACT

Anti-doping authorities have high expectations of the *athlete steroidal passport* (ASP) for anabolic–androgenic steroids misuse detection. However, it is still limited to the monitoring of known well-established compounds and might greatly benefit from the discovery of new relevant biomarkers candidates. In this context, steroidomics opens the way to the untargeted simultaneous evaluation of a high number of compounds. Analytical platforms associating the performance of ultra-high pressure liquid chromatography (UHPLC) and the high mass-resolving power of quadrupole time-of-flight (QTOF) mass spectrometers are particularly adapted for such purpose. An untargeted steroidomic approach was proposed to analyse urine samples from a clinical trial for the discovery of relevant biomarkers of testosterone undecanoate oral intake.

Automatic peak detection was performed and a filter of reference steroid metabolites mass-to-charge ratio (m/z) values was applied to the raw data to ensure the selection of a subset of steroid-related features. Chemometric tools were applied for the filtering and the analysis of UHPLC-QTOF-MS^E data. Time kinetics could be assessed with N-way projections to latent structures discriminant analysis (N-PLS-DA) and a detection window was confirmed. Orthogonal projections to latent structures discriminant analysis (O-PLS-DA) classification models were evaluated in a second step to assess the predictive power of both known metabolites and unknown compounds. A shared and unique structure plot (SUS-plot) analysis was performed to select the most promising unknown candidates and receiver operating characteristic (ROC) curves were computed to assess specificity criteria applied in routine doping control. This approach underlined the pertinence to monitor both glucuronide and sulphate steroid conjugates and include them in the *athletes passport*, while promising biomarkers were also highlighted.

1. Introduction

While traditional anti-doping protocols rely on the direct detection of prohibited substances or their metabolites, strong interest has developed recently in the longitudinal monitoring of parameters to examine direct as well as indirect markers over time. This approach is particularly useful in the case of endogenous compounds that cannot be differentiated easily from synthetic analogues. The introduction of the athlete biological passport [1] constitutes a major breakthrough in the conception of doping control. It aims at the indirect assessment of doping substance abuse by the detection of unexpected patterns related to a choice of selected biological variables. Control strategies progressively evolve from the conventional comparison of measured values with population references and tolerance limits to more sophisticated methodologies integrating the evaluation of idiosyncratic levels, variability and individual limits. In this viewpoint, each individual constitutes hence a self-reference. Despite being originally designed to assess parameters related to blood doping

☆ This paper is part of the special issue entitled: Fight Against Doping in 2011, Guest-edited by Neil Robinson (Managing Guest Editor), Martial Saugy, Patrice Mangin, Jean-Luc Veuthey, Serge Rudaz and Jiri Dvorak.

* Corresponding author at: School of Pharmaceutical Sciences, University of Geneva, 20 Bd d'Yvoy, 1211 Geneva 4, Switzerland. Tel.: +41 22 379 65 72; fax: +41 22 379 68 08.
E-mail address: serge.rudaz@unige.ch (S. Rudaz).

[1] Both authors have contributed equally to this work.

0379-0738/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2011.07.023

(athlete haematological passport [2]), such an approach is likely to face other doping practices. For instance, the athlete steroidal passport (ASP) [3] is a very promising tool to detect the misuse of anabolic–androgenic steroids (AAS).

AAS constitute a class of hormonal compounds related to testosterone, its metabolites and precursors. As modified levels of several compounds have been highlighted after steroid administration [4], a series of these diagnostic markers which are close relatives of testosterone were included in the so-called *steroid profile* for monitoring purpose. This profile comprises testosterone itself (T), epitestosterone (E), androsterone (A), etiocholanolone (Etio), dehydroepiandrosterone (DHEA), 5α- or 5β-androstane-3α,17β-diol (5α- and 5β-Adiol) and dihydrotestosterone (DHT). Reference ranges and ratio were empirically estimated from clinical trials and threshold values were derived from the results to constitute decision rules applicable to screening procedures in routine control [5]. As an example, the most representative parameter regarding AAS misuse detection is the T/E $\geq$ 4.0'' rule [6]. The unambiguous distinction between compounds naturally present in the human body and exogenous intake of synthetic AAS relies however on $\delta^{13}C$ values measured by isotope ratio mass spectrometry (IRMS) [7]. Therefore, the isotopic signature of suspicious samples failing to meet the decision rules derived from the *steroid profile* is analysed with IRMS for a definitive confirmation.

Nonetheless, the preliminary screening of AAS misuse still constitutes a major concern for anti-doping authorities. AAS are excreted in urine as phase II metabolites after glucuro- or sulpho-conjugation. This process intends to make molecules more polar and facilitate their excretion. A large heterogeneity related to genetic factors, *i.e.* UDP-glucuronosyltransferase 2B17 (UGT2B17) enzyme polymorphisms [8–10], underlined the limitations of traditional population-based criteria in screening protocols. Therefore, the longitudinal monitoring of the conventional *steroid profile* with the ASP is expected to greatly enhance the ability to detect AAS doping, but improvements might be still available based on relevant biomarkers discovery. Additionally, while glucuronide conjugates constitute well-established markers, the examination of sulphate metabolites is expected to provide complementary indications regarding steroid excretion and metabolism [11,12].

A deconjugation step is usually applied in routine screening prior to sample analysis by gas chromatography–mass spectrometry (GC–MS) [13]. Although these protocols produce highly reliable results, they are time-consuming and fail to provide a comprehensive monitoring of phase II metabolites. In this context, the coupling of liquid chromatography to mass spectrometry (LC–MS) was demonstrated as a valuable analytical alternative for the direct analysis of intact conjugated steroids [14,15] and presents the potential to measure a wide range of metabolites, including glucuronide and sulphate conjugates with high sensitivity and selectivity. In particular, ultra-high pressure LC (UHPLC) hyphenated to high mass-resolving power time-of-flight (TOF) or quadrupole TOF (QTOF) mass spectrometers was reported to be particularly well suited for the analysis of steroid metabolites in biofluids such as urine [16]. Such analytical platforms are able to perform a full mass range acquisition with high accuracy and sensitivity, and QTOF-MS provides additional information by allowing exact mass measurements of both (de)protonated molecules and fragment ions. QTOF-MS^E involves the simultaneous acquisition of both low and high energy collision products. Parent and fragment mass information are thereby assessed in a single run. This ability is of major importance as it offers structural information required for the identification of unknown biomarkers in the context of untargeted analyses.

The recent profusion of information provided by holistic approaches (''omics'') allows innovative perspectives. Untargeted analyses constitute hence a potent tool for the discovery of biomarkers related to a physiological response and the diagnosis of complex phenotypes, with the help of chemometric methods [17]. Metabolomics in general and steroidomics in particular represent therefore promising strategies in the context of doping analysis for the monitoring of changes in levels and ratio of endogenous steroids. Studies assessing differences between urinary signature of animals after anabolic agents treatment for doping purpose [18], or applied to agricultural meat production [19,20] have been reported over the last past years.

The present work proposes an untargeted steroidomic approach for the analysis of urine samples from a clinical trial. It pertains to the discovery of new relevant biomarkers for the detection of testosterone intake. Chemometric tools were applied for the filtering and the analysis of UHPLC-QTOF-MS^E data. After proper validation and identification, the most relevant candidates could constitute appropriate parameters to include in the ASP.

2. Experimental

2.1. Urine samples

Samples were obtained from a previous clinical trial carried out over one month and designed to assess the effects of multiple oral doses of testosterone undecanoate administration in healthy male subjects. A thorough description of the study, simulating common doping practices, can be found elsewhere [21]. Two groups of volunteers were selected, namely the placebo (PLACEBO, $n = 9$) and the testosterone (TESTO, $n = 8$) groups. Pills containing 300 mg of mannitol for the placebo group, and 80 mg of testosterone undecanoate and 115 mg of mannitol for the testosterone group were administered.

Six urine samples were taken from each volunteer. A first sample was taken before the beginning of the study (t00). A urinary excretion monitoring over 24 h was measured on the 24th day of the study, 0, 4, 8 and 24 h after pill intake, corresponding to the samples t01, t02, t03 and t04, respectively. Finally, a last urine sample was taken on the 38th day of the treatment, two weeks after the last pill intake. Two samples could not be obtained at t01 (one in each group), due to technical problems. This selection of samples intends to provide insights into the metabolic kinetics of steroids in urine after testosterone undecanoate oral administration by comparing a basal state (t00) before any pill intake, a time course after intake (t01–t04) and a presumed return to the initial situation, *i.e.* resilience (t05), 200 h after the last pill intake. A total of 100 samples were thus gathered. As discussed by Baume et al. [21,22], volunteers U03 and U25 were shown to present the UGT2B17 del/del homozygote enzyme polymorphism. Indeed, it was observed that the U03 T/E ratio never exceeded 0.5, even if IRMS analysis confirmed an exogenous testosterone intake. Additional genotyping pointed out a deletion of this gene for the volunteers U03 and U25, as reported elsewhere [23]. Moreover, volunteer U11 was identified as a high-rate excretor and volunteer U14 was demonstrated to possess a massive steroid metabolism.

2.2. Analysis of urine samples

An analytical protocol based on a UHPLC-QTOF-MS^E platform was used for the quantification of steroid metabolites in urine. A thorough description of the method can be found in [16]. Two functions were acquired in parallel in the QTOF-MS^E mode. It allowed the simultaneous acquisition of m/z over the entire mass range (m/z 95–1000) in a first function at low collision energy (5 eV), while a second function providing rich fragmentation patterns was obtained with ramped collision energy (from 5 to 70 eV). Such analytical methodology opens the way to an extended monitoring of a high number of conjugated steroid metabolites excreted in urine and provides precious structural information for the identification of new chemical entities. All urine samples were therefore analysed following this protocol for steroidomic analysis.

The quality of data generated during the entire analytical process was assessed by the use of quality control samples, allowing the determination of retention time drifts and m/z precision [24]. Samples were randomised and analysed in triplicate to avoid batch effects [16]. Moreover, isotopically labelled standards were used for normalisation.

2.3. Data pretreatment

Untargeted raw data pre-treatment was performed with the MarkerLynx™ software. It included baseline correction, peak detection, alignment and normalisation. Peak detection tolerance parameters of 0.05 Da for m/z values and 30 s for retention time shifts were set. Ions meeting these criteria in the different samples were considered as the same variable. A series of 5750 variables characterising each sample was obtained over the selected mass range (m/z 95–1000). Each variable corresponded therefore to a peak area measured in all samples for a given m/z signal measured at a specific chromatographic retention time. The ions selection was then

performed. A list of reference m/z values related to endogenous steroid conjugates (glucuronidated and sulphated forms) belonging to the *sterol lipids* class of the Lipid Maps database [25,26] was used. Each of the 5750 ions was compared with the 104 reference m/z values of the list with a tolerance interval of ±0.025 Da. A dataset of 100 urine samples and 234 putative steroid-related variables corresponding to a chromatographic peak in the first function acquired at low collision energy was obtained. The supposed conjugated steroid structure of each variable was assessed by the monitoring of its fragmentation pattern provided by the energy ramp (from 5 to 70 eV) function.

Additionally, a series of 10 targeted known steroid metabolites was independently quantified. It included both glucuronidated and sulphated forms of steroids related to the *steroid profile*, *i.e.* testosterone glucuronide (TG), testosterone sulphate (TS), epitestosterone glucuronide (EG), epitestosterone sulphate (ES), androsterone glucuronide (AG), androsterone sulphate (AS), etiocholanolone glucuronide (EtioG), etiocholanolone sulphate (EtioS), dehydroepiandrosterone glucuronide (DHEAG) and dehydroepiandrosterone sulphate (DHEAS). Each targeted metabolite was identified and confirmed by its retention time and accurate mass in comparison with pure standards injected separately [16].

2.4. Data structure

The temporal follow-up of several biomarkers in a series of individuals generates data which cannot be naturally summarised in a conventional data table. Since a one-dimensional vector of variables is obtained by measuring compounds in each urine sample, a two-dimensional structure (P variables × Q time points) characterises the concatenated sample vectors in a table for each individual. When considering the whole dataset, a cube of data (N individuals × P variables × Q time points) is obtained by gathering the data tables from all volunteers in a three-dimensional data structure, *i.e.* a three-way tensor. As a result, classical chemometric tools such as PCA and PLS cannot be applied unless prior matricisation. Matricisation (also called unfolding or flattening) intends to reorganise the elements of a tensor to form a two-dimensional matrix. Different matricisation alternatives are offered to reorder a tensor, as frontal, horizontal or lateral slices from a three-way data cube can be rearranged by a row-wise or column-wise concatenation (see Fig. 1).

Multi-way chemometric approaches are available to deal with such high-dimensional data structures. These techniques are well-suited for the analysis of metabolic profiles with intrinsically multicollinear variables and temporal monitoring by taking advantage of all the information provided by the whole data structure [27]. Multi-way algorithms such as PARAFAC [28] and Tucker [29] provides unsupervised models, while N-way projections to latent structures (N-PLS) is an extension of the supervised two-way PLS decomposition of conventional two-way data table to higher orders arrays [30]. The latter provides an attractive approach to build predictive models on raw data structures and captures efficiently the underlying structure of the data.

2.5. Data analysis

N-PLS discriminant analysis (N-PLS-DA) models and receiver operating characteristic (ROC) curves evaluations were performed under the MATLAB® 7 environment (The MathWorks, Natick, USA). N-PLS-DA was assessed with routines implemented in the N-Way Toolbox (version 3.1) [31]. O-PLS discriminant analysis (OPLS-DA) models were evaluated with the SIMCA-P software (version 12, Umetrics, Umeå, Sweden). For each model, a leave-one-subject-out cross-validation was performed to assess the model fit. Indices of accuracy, sensitivity and specificity were computed to measure classification performance. Model validity was verified using permutation tests (Y-scrambling) and CV-ANOVA [32].

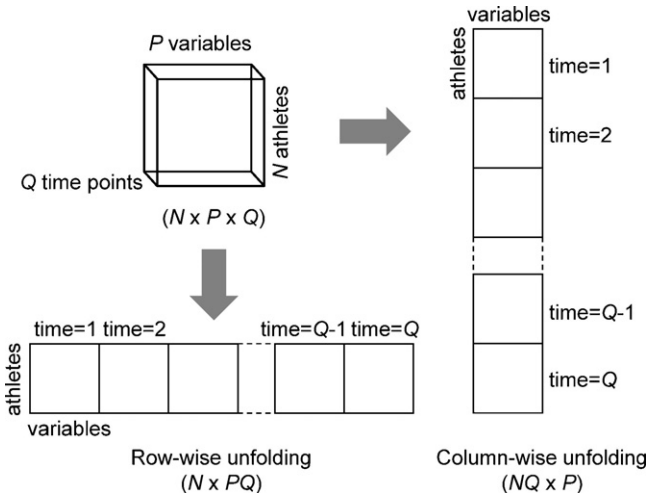


Fig. 1. Three-way tensor matricisation with row-wise and column-wise unfolding.

3. Results and discussion

3.1. Targeted profile

Ten compounds including glucuronidated and sulphated forms of endogenous steroids were quantified using the targeted method described in Badoud et al. [16]. 5α-Androstane-3β,17β-diol-3-glucosiduronate was not evaluated because of its low concentration range close to the limit of detection, leading to irrelevant patterns. A tri-dimensional data structure including 17 individuals described by 10 steroid conjugates measured at six time points ($17 \times 10 \times 6$) was initially obtained. Data were centred across the sample and the variable mode, while unit variance scaling was applied within the variable mode. Scaling to unit variance within the variable mode is required to assign the same importance to the variability of each of the steroid conjugates, regardless of its signal intensity. A discriminant analysis was performed by building a N-PLS model with a dependent class variable (PLACEBO vs. TESTO class). Two latent variables were computed to obtain the best prediction ability as estimated by leave-one-out cross-validation. The following performance indices were obtained: accuracy 76.5% (overall rate of correct predictions), sensitivity 62.5% (predicted TESTO/true TESTO) and specificity 88.9% (predicted PLACEBO/true PLACEBO). Scatter plots (LV1 vs. LV2) depicting the sample mode (A), the variable mode (B) and the time mode (C) of the N-PLS-DA model are presented in Fig. 2.

The sample mode (Fig. 2A) indicated a separation between the PLACEBO and TESTO groups. However, two volunteers from the TESTO group exhibited an unexpected *steroid profile*, namely U03 and U11. These individuals were either previously genotyped as del/del homozygote for the UGT2B17 gene (U03) [23] or detected as a high-rate steroid excretor (U11), in accordance with the biological variability previously observed.

Different clusters of variables were observed in the N-PLS variables mode (Fig. 2B). A group of variables including AS, AG, EtioS, EtioG and TG was associated with the TESTO class, indicating increased levels in this class, while ES and EG were related to the PLACEBO class. As expected, DHEAG had no clear influence, as it was located at the centre of the projection plane, reflecting no concentration alteration due to testosterone intake. No conclusion regarding TS and DHEAS could be derived from the model probably due to the high variability observed. These results are also in accordance with previous studies and anti-doping protocols since the *steroid profile* was chosen to monitor targeted metabolites as indicators of steroid misuse [33]. While AG, EtioG and TG are known to be increased after testosterone intake, EG and ES concentrations are less affected and even decreased after high testosterone doses [34].

The time mode (Fig. 2C) was of high interest, as it provided a representation of the time-related variation pattern in a two-dimensional subspace. A trajectory could be drawn by connecting the time points chronologically and relevant information regarding the time kinetics of testosterone excretion could be deduced. Starting from a basal situation t00, only few changes were observed at t01 on the basis of the ten variables. The t01 time point corresponded to 0 h after the pill intake at the 24th day of the clinical trial. Small differences from the basal situation could be related to changes in the *steroid profile* due to prior intake of testosterone undecanoate (3 pills a week for 3 weeks [21]). A detection window was evidenced between t02 and t03, pointing out clear alterations of steroid concentrations. Then the t04 and t05 time points indicated a progressive return to the basal situation. These results confirmed the relevance of the *steroid profile* for a

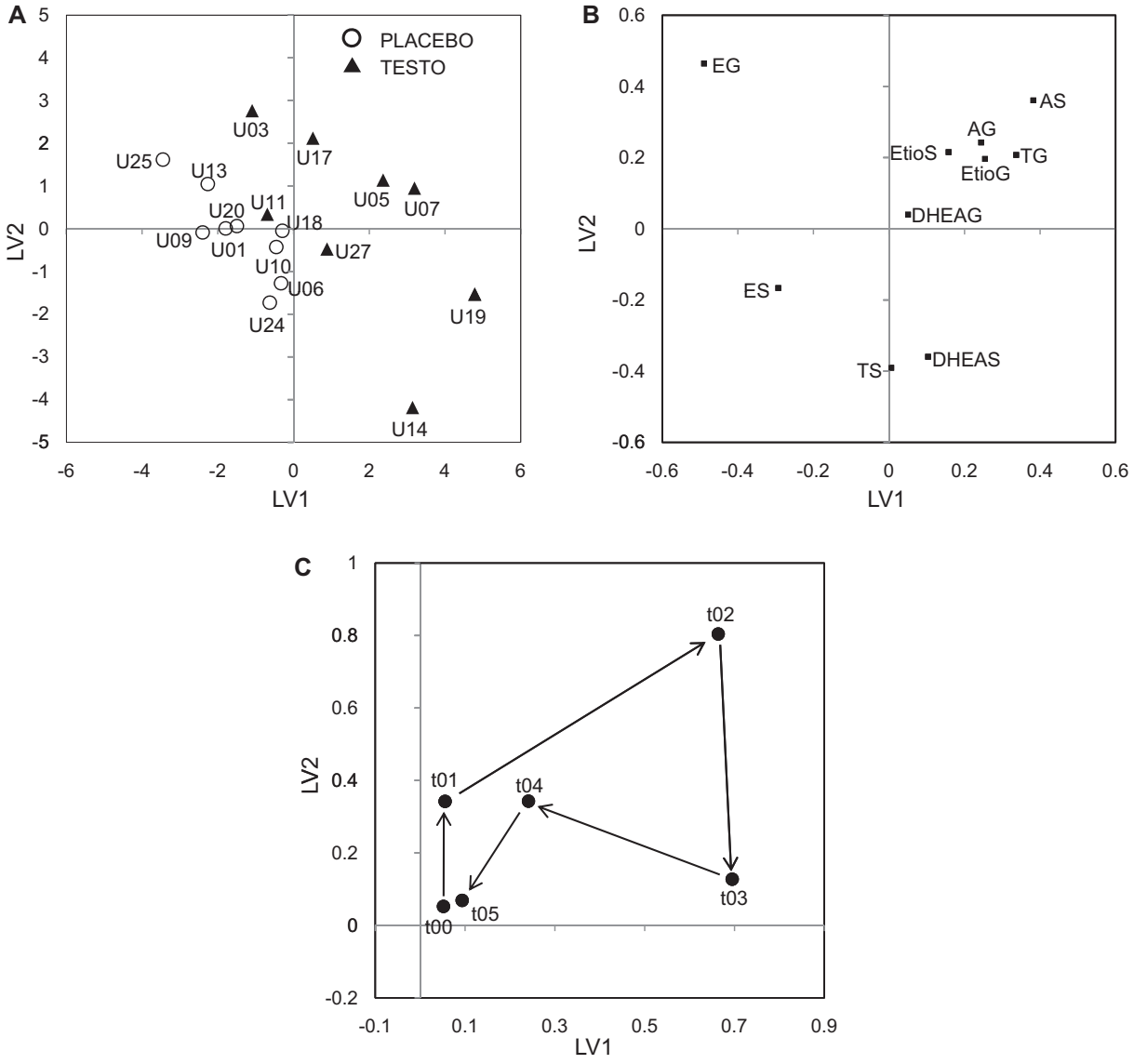


Fig. 2. Targeted profile N-PLS-DA scatter plots (LV1 vs. LV2) with respect to (A) the sample mode, (B) the variable mode and (C) the time mode. Volunteers from the PLACEBO group are symbolised by white circles (○) and individuals from the TESTO group by black triangles (▲). Variables are symbolised by black squares (■) and time points by black circles (●).

small detection window corresponding to the marked phase of testosterone metabolites excretion. There is therefore a great interest to increase the number of biomarkers of testosterone intake to ensure a comprehensive monitoring of the urinary steroid content and extend the detection window of steroid misuse [35,36].

3.2. Extended profile

An additional analysis of the data was then envisaged to explore the wealth of information provided by an extended steroid monitoring with 234 steroid-related variables. A similar pre-processing was applied, *i.e.* centring across the sample and the variable mode and unit variance scaling within the variable mode. A N-PLS-DA model with two latent variables was obtained by leave-one-out cross-validation. The following performance indices were computed: accuracy 82.4% (overall rate of correct predictions), sensitivity 75% (predicted TESTO/true TESTO) and specificity 88.9% (predicted PLACEBO/true PLACEBO). It has to be noticed that an increased sensitivity was observed with the extended profile, without deteriorating the specificity and a clear separation of the two classes was noticeable on the sample mode scatter plot (Fig. 3A). A compact cluster of individuals was obtained for the TESTO class. Individuals from the PLACEBO class exhibited a looser grouping with a score close to zero on LV1 for volunteer U10 who was already located at the centre of the projection plane in the targeted profile modelling (see Fig. 2A). U03 and U11 were confirmed as the most difficult individuals of the TESTO class to classify. These three volunteers (U10, U03 and U11) constituted the only misclassified cases according to their cross-validated predictions. Substantially similar trends were observed in the projection scores of the individuals when compared to the targeted profile analysis. The results obtained with an extended steroid monitoring were therefore in accordance.

Additional information about kinetics of steroid metabolites excretion could be inferred by the trajectory in the time mode (see Fig. 3C). Starting again from a basal situation t00, marked changes were observed at t01 allowing an earlier detection of testosterone intake compared to the targeted analysis, evolving into a marked phase at t02 and t03. Time kinetics indicated then a return to the

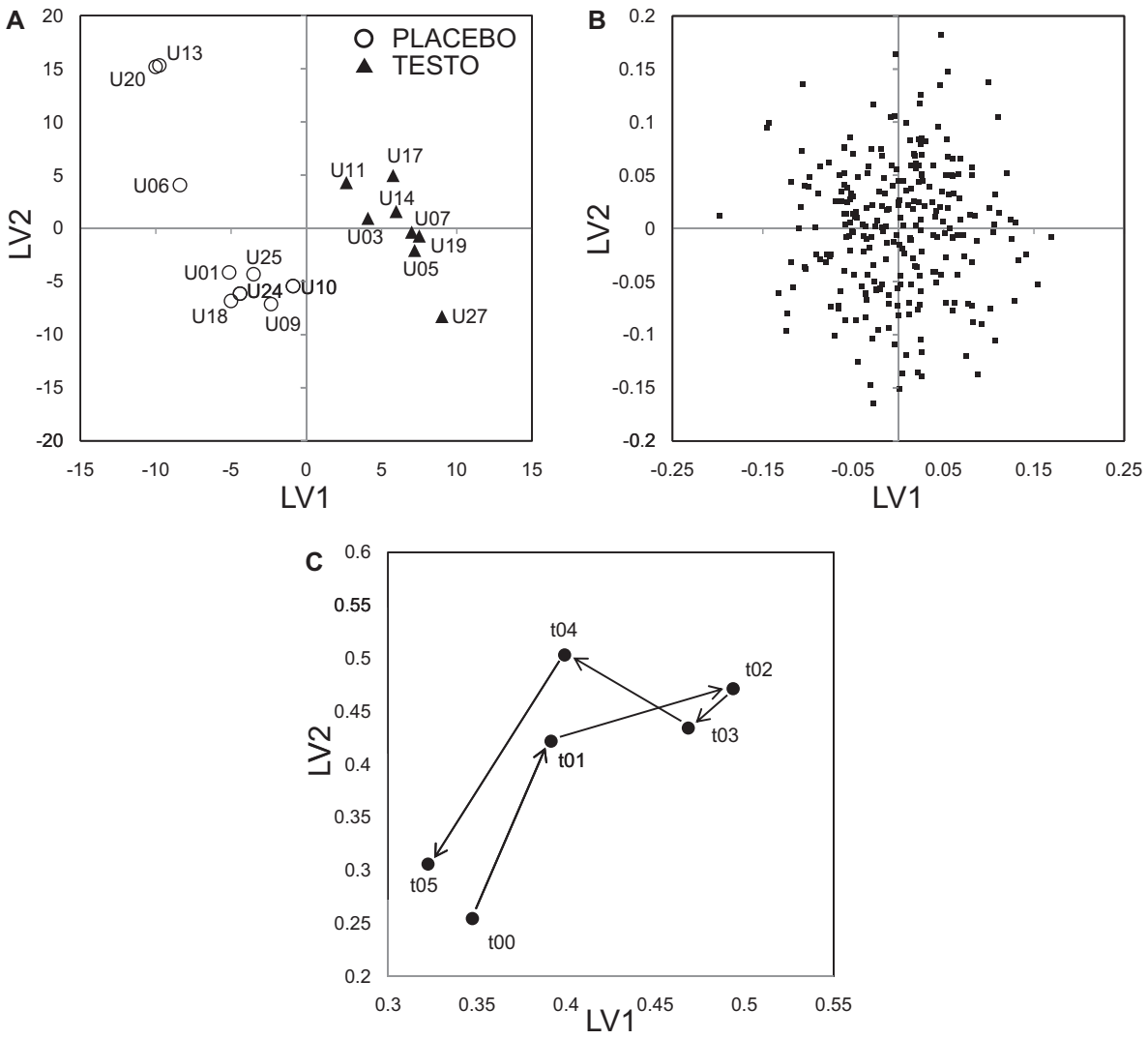
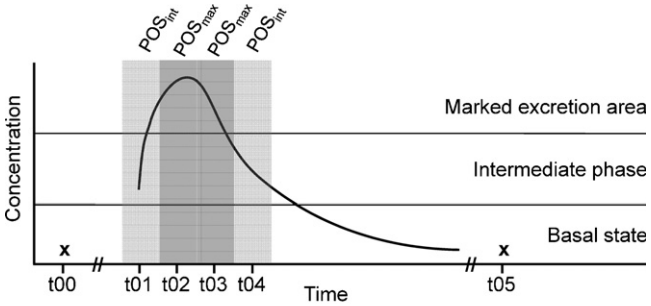


Fig. 3. Extended profile N-PLS-DA scatter plots (LV1 vs. LV2) with respect to (A) the sample mode, (B) the variable mode and (C) the time mode. Volunteers from the PLACEBO group are symbolised by white circles (○) and individuals from the TESTO group by black triangles (▲). Variables are symbolised by black squares (■) and time points by black circles (●).

moderate state at t04 and to the basal state at t05. These results suggest the opportunity to find biomarkers of testosterone misuse related to steroid structures and enlarge its detection window in urine based on the 234 steroid-related variables of the extended profile.

3.3. O-PLS-DA modelling and SUS-plot analysis

Despite revealing noteworthy information about excretion kinetics, a rigorous monitoring of known time points was necessary to characterise each individual in both N-PLS-DA models. Actually, these models were not tailored to assess single urinary samples lacking temporal information. This aptitude remains however a major concern in the context of doping control. Indeed, the longitudinal monitoring introduced in the *athlete biological passport* depends only on the sequence of measurements, even if particular care is required when considering measurements separated by less than five days [1]. A column-wise unfolding was therefore performed and the temporal information gathered during the first step of multiway modelling was used to reduce the detrimental effects of matricisation by defining three new classes of samples. A cyclic excretion pattern putatively divided in three phases, *i.e.* a basal state (t00 and t05), two intermediate situations (t01 and t04) and a phase of marked excretion (t02 and t03), was observed from the trajectory in the time mode scatter plot after testosterone administration (Figs. 2C and 3C). Consequently, a negative class (NEG) was defined to include all samples from the PLACEBO group and samples t00 and t05 from the TESTO group. A class corresponding to the



GROUP	t00	t01	t02	t03	t04	t05
PLACEBO	NEG	NEG	NEG	NEG	NEG	NEG
TESTOSTERONE	NEG	POS_{int}	POS_{max}	POS_{max}	POS_{int}	NEG

Fig. 4. Theoretical excretion kinetics and class distribution.

intermediate situations was named POS_{int} and included all samples from the TESTO group at t01 and t04. A class related to the marked excretion phase was named POS_{max} and included all samples from the TESTO group at t02 and t03. Theoretical excretion kinetics and class distribution of the urine samples are described in Fig. 4.

The contribution of both the targeted *steroid profile* involving 10 known compounds of biological relevance and the extended *steroid profile* including 234 unidentified putative steroid metabolites, extracted from the first MS function, was evaluated by the concatenation of both data tables. Autoscaling was applied to the fused table containing all data (234 + 10 variables) to assign the same importance to the variability of each analyte, without considering its signal intensity. As urine samples taken from the same volunteer are not independent from a statistical point of view, a leave-one-subject-out cross-validation was performed. This procedure is based on classification models built with all data but leaving all the samples from a given subject out. It is repeated once per subject to ensure the robustness of the model and its generalisation ability to new samples. Accuracy, sensitivity and specificity indices were computed to measure classification performance.

Two independent O-PLS-DA models were evaluated to distinguish steroid patterns from the NEG vs. POS_{int} and the NEG vs. POS_{max} comparison, respectively. The O-PLS method is an extension of the PLS algorithm designed to summarise the predictive part of the data in the first latent variable, *i.e.* the predictive component [37]. Within-class variation that is not related (orthogonal) to the class response is contained in the so-called orthogonal component(s) [38]. This partition of the initial variation allows an increased interpretability of the discriminant models in terms of variable contributions and biological relevance.

Both models were statistically significant (CV-ANOVA p-value < 0.05) and were composed of one predictive and one orthogonal component. The following performance indicators were computed:

(a) NEG vs. POS_{int} model: accuracy 89.3%, sensitivity 60.0% (predicted POS_{int}/true POS_{int}) and specificity 95.7% (predicted NEG/true NEG).
(b) NEG vs. POS_{max} model: accuracy 88.2%, sensitivity 68.8% (predicted POS_{max}/true POS_{max}) and specificity 92.8% (predicted NEG/true NEG).

In both models, a bias toward high specificity was probably due to an unbalanced dataset with a majority of NEG samples (~80%). Fortunately, this bias is not an issue in the context of doping analysis. It could even become desirable since false positives should be avoided, although the screening procedure is always followed by a confirmation step by IRMS in case of a presumptive analytical finding. High sensitivity is needed to efficiently detect suspicious samples but a high screening specificity intends to reduce the number of subsequent time-consuming experiments.

The correlation vector $corr(t_p,X)$ from the predictive component (first latent variable) of each model was extracted to build a shared and unique structure plot (SUS-plot). This representation intends to evaluate the contributions of variables to different classification models by the comparison with a common reference [39], *i.e.* the NEG class in that case. The scores plots of both O-PLS-DA models and the corresponding SUS-plot are presented in Fig. 5A, B and C, respectively.

The variable contributions were consequently analysed with respect to their position on the SUS-plot. The upper right region of the plot displayed in Fig. 5D was of particular interest as it included variables sharing increased signals in both the POS_{int} and POS_{max} classes, compared to urine samples from the NEG class. While ES and EG had almost no influence to distinguish the three classes, a contribution of the remaining 8 known steroid metabolites toward both the class related to the marked phase of excretion and the intermediate situation was highlighted. These results are consistent with established doping control protocols based on parameters such as the T/E ratio. As two independent O-PLS-DA models were computed, small changes related to either the intermediate situation (POS_{int}, *i.e.* t01 and t04) or the marked phase (POS_{max}, *i.e.* t02 and t03) of the testosterone metabolism could be detected separately. A selection of the most promising unknown compounds was made according to the SUS-plot (Fig. 5D) and a restrained list of 12 original candidates was obtained. Some of them corresponded to known steroid metabolites, such as compound #01 (m/z 463.2350 at 11.49 min) that was identified as TG thanks to the comparison of its retention time and exact mass with the analysis of the pure standard. Moreover, 5α-AdiolG and 5β-AdiolG were highlighted as two major peaks of the m/z 467.2649 extracted ion chromatogram at 16.46 (compound #11) and 17.44 min (compound #8), respectively. They were eluted in the chromatogram before AG and EtioG, pointing out their higher polarity. Moreover, the ion at m/z 467.2649 corresponded to the theoretical exact mass of $[M-H]^-$ 5α- and 5β-AdiolG, which was 467.2645, with an error of 0.8 ppm. The structure could also be attested by the low and high ramped energy spectra obtained in the MS^E mode. Indeed, as illustrated in Fig. 6, the deprotonated molecular ion of 5α-AdiolG (m/z 467.2649) was present in the first function acquired at low collision energy (5 eV), while its fragments could be observed in the second function. The fragments corresponded to those typically arising from the glucuronide itself at m/z 157.0176, m/z 113.0298, m/z 85.0379 and m/z 75.0187, as reported elsewhere [40]. Additionally, neutral loss of glucuronic acid (loss of m/z 176.0312) could be observed with the presence of m/z 291.2327 corresponding to the aglycone part. The analysis of elemental composition was performed on this ion and the formula $C_{19}H_{31}O_2$ for the $[M-H]^-$ deprotonated molecule was attributed with an error of less than 10 ppm to the deconjugated AdiolG molecule. Although no fragment could be attributed to the aglycone steroid itself, those metabolites could be attributed to the class of glucuroconjugated steroids.

3.4. Receiver operating characteristic (ROC) curve analysis

The traces of the most promising biomarkers were extracted from the first MS function and checked to ensure a correct peak assignment and an accurate integration of the chromatographic peaks. A normalisation was performed by correcting each peak area using the closely eluted and structurally nearest isotopically labelled standard. Peaks # 08, 10, 11 and 12 were corrected with the area of TG-D3; peaks # 04 and 09 were normalised according to AS-D4, peaks # 02 and 03 were normalised by ES-D3, while # 07 was corrected with EG-D3. Finally peak areas of compounds # 05 and 06 were normalised with respect to AG-D4. The diagnostic performance of both 10 known metabolites and 11 remaining unknown candidates was then evaluated using ROC analysis. In order to circumvent any temporal constraint, POS_{int} and POS_{max} urine samples were gathered in a time-independent POS class and compared with the NEG group.

ROC curves, depicted in Fig. 7 for TG, TG/EG, AS, compounds #4, #05 and DHTG (#6), are powerful indicators of the accuracy of a test to discriminate between two populations in the case of overlapped value distributions. For every possible threshold value, performance indices are computed and reported graphically. Each point of the ROC curve corresponds hence to a sensitivity/specificity couple associated with a particular decision threshold. The area under the curve (AUC) is a global index frequently used to summarise the information contained in the ROC curve. However, misleading conclusions can arise when

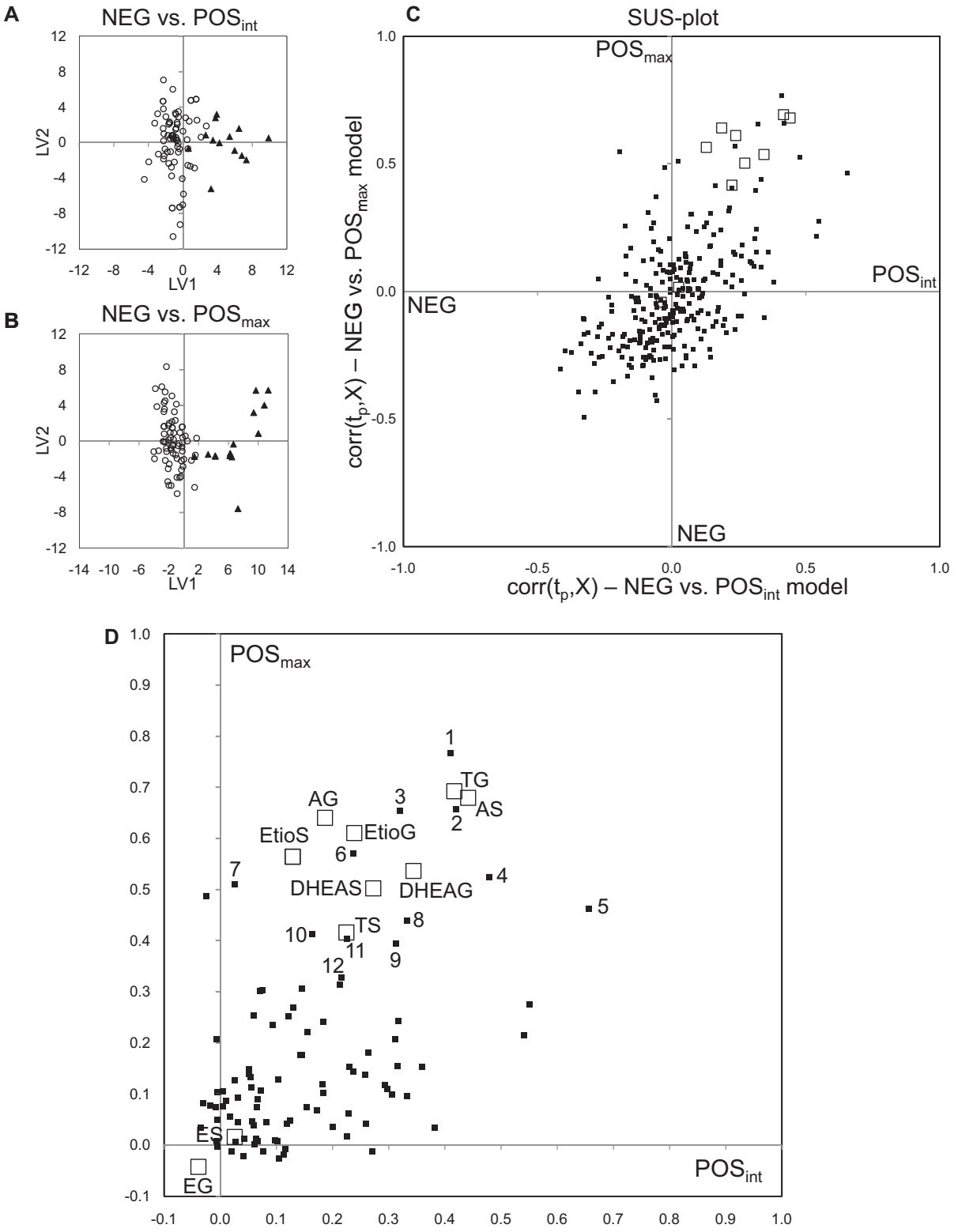


Fig. 5. O-PLS-DA models between (A) the NEG vs. POS_{int} class and (B) the NEG vs. POS_{max} class. Volunteers from the PLACEBO group are symbolised by white circles (○) and volunteers from the TESTO group by black triangles (▲). (C) SUS-plot reflecting the variables' contribution to the predictive component of both O-PLS-DA models. (D) SUS-plot variables Map. Known metabolites (*i.e.* from the *steroid profile*) are symbolised by white squares (□) and unknown variables by black squares (■).

interest resides in a specific region of the curve and not in the entire range. As mentioned above, this is particularly the case in doping control where false positive cases must be avoided, since extensive confirmation procedures are needed. High specificity requirements must therefore be fulfilled to validate a potential biomarker for screening purpose and the area in the specificity range 70–100% ($AUC_{>70}$) was computed. It has to be noted that total area values (AUC_{total}) were assessed for comprehensiveness but were not further discussed. Besides, sensitivity was computed for all candidates at a fixed specificity level of 99% ($SENSI_{SPEC99\%}$). These results are summarised in Table 1.

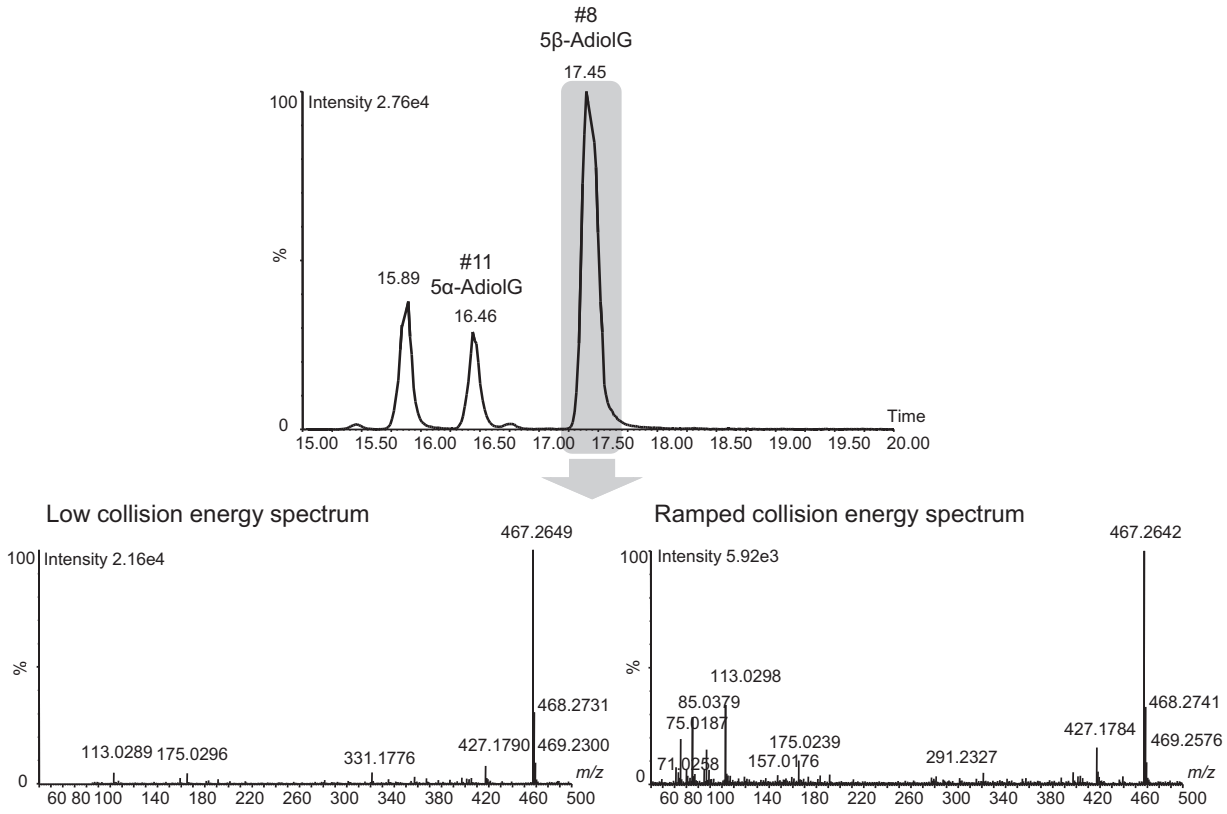


Fig. 6. Low (5 eV) and ramped (from 5 to 70 eV) collision energy MS^E spectra from the extracted trace at m/z 467.2649 (volunteer U07 at t03).

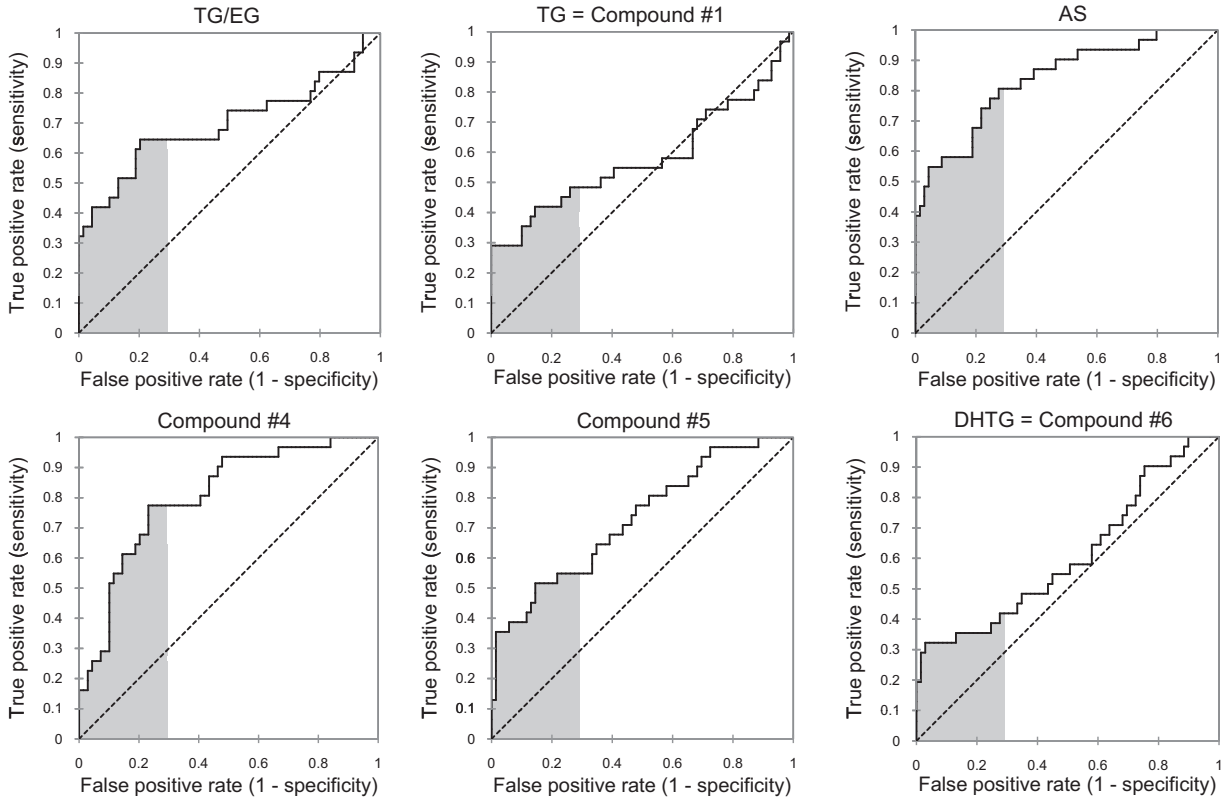


Fig. 7. ROC curves representing the true positive rate (sensitivity) vs. the false positive rate (1 − specificity) according to the cut-off limit for concentrations of TG/EG, TG, AS, Compound #4, Compound #5 and DHTG. The $AUC_{>70}$ region is indicated in dark grey.

Table 1
ROC analysis results. AUC_{total}: ROC area in the specificity range 0–100%; $AUC_{>70}$: ROC area in the specificity range 70–100%; $SENSI_{SPEC99\%}$: sensitivity measured at a fixed specificity level of 99%.

Compound	*m/z*	Ret. time	AUC_{total}	$AUC_{>70}$	$SENSI_{SPEC99\%}$	Putative structure
TG/EG	–	–	0.701	0.135	35%	–
TG	463.2332	11.52	0.659	0.116	33%	–
TS	367.1736	12.64	0.608	0.085	19%	–
EtioG	465.2488	20.54	0.673	0.118	22%	–
EtioS	369.1736	20.34	0.646	0.119	26%	–
AG	465.2488	21.18	0.652	0.125	22%	–
AS	369.1736	19.57	0.837	0.185	39%	–
DHEAG	463.2332	13.60	0.485	0.063	0%	–
DHEAS	367.1736	16.40	0.720	0.125	7%	–
EG	463.2332	16.88	0.461	0.040	0%	–
ES	367.1736	13.68	0.544	0.067	0%	–
#1	463.2350	11.49	0.659	0.116	33%	TG
#2	383.1539	11.40	0.584	0.085	19%	
#3	383.1562	15.20	0.499	0.060	3%	
#4	369.1752	18.46	0.806	0.155	16%	AS, EtioS or DHTS isomer
#5	481.2446	12.89	0.728	0.132	35%	Hydroxylated form of AG, EtioG or DHTG
#6	465.2483	15.83	0.608	0.099	29%	DHTG
#7	463.2345	18.45	0.586	0.078	23%	
#8	467.2649	17.44	0.618	0.091	3%	5β-AdiolG
#9	448.3064	22.94	0.586	0.092	10%	
#10	477.2073	8.86	0.539	0.060	0%	
#11	467.2657	16.46	0.592	0.064	6%	5α-AdiolG
#12	481.2451	12.47	0.530	0.056	0%	

First, the TG/EG ratio was computed to assess its performance with respect to the clinical trial at hand. $AUC_{>70}$ and $SENSI_{SPEC99\%}$ values of 0.125 and 35% respectively underlined the intrinsic difficulties related to the cohort under consideration, regarding its biological variability and genetic polymorphisms. Among established known compounds, TG considered alone was confirmed as a relevant biomarker with $AUC_{>70}$ = 0.116 and $SENSI_{SPEC99\%}$ = 33%. TS concentrations were probably too close to the limit of detection to provide sufficiently reliable values and ensure classification criteria ($AUC_{>70}$ = 0.085 and $SENSI_{SPEC99\%}$ = 19%). EtioG ($AUC_{>70}$ = 0.118, $SENSI_{SPEC99\%}$ = 22%) and EtioS ($AUC_{>70}$ = 0.119, $SENSI_{SPEC99\%}$ = 26%) presented similar results, highlighting their pertinence for testosterone misuse control.

Interestingly, the performance indices obtained for AS ($AUC_{>70}$ = 0.185, $SENSI_{SPEC99\%}$ = 39%) were notably higher than for AG ($AUC_{>70}$ = 0.125, $SENSI_{SPEC99\%}$ = 22%). It has to be noted that AS was already highlighted as a relevant parameter to detect DHEA administration [11]. These values suggested that sulphates metabolism probably supplied additional valuable information and underlined the relevance to monitor sulphated forms of steroid metabolites to extend the detection window related to the testosterone intake. As expected, DHEAG ($AUC_{>70}$ = 0.063 and $SENSI_{SPEC99\%}$ = 0%), EG ($AUC_{>70}$ = 0.040 and $SENSI_{SPEC99\%}$ = 0%) and ES ($AUC_{>70}$ = 0.067 and $SENSI_{SPEC99\%}$ = 0%) revealed no detection ability. Indeed epitestosterone is known to be stable after low testosterone administration, while it decreases with high doses administered [34]. It is used as internal reference when assessing the T/E ratio. These results are in accordance with recent studies [36]. On the other hand, DHEAS exhibited an interesting $AUC_{>70}$ value of 0.125, but its $SENSI_{SPEC99\%}$ performance index was only of 7%. A high variability and intensity normalisation issues may be responsible for these results, as no deuterated standard was commercially available for that particular compound.

Among unknown compounds, three candidates revealed very interesting performance indices. Compound #04 exhibited a high $AUC_{>70}$ value of 0.155 but a rather modest $SENSI_{SPEC99\%}$ index of 16%. Inversely, compound #06 revealed an $AUC_{>70}$ value of 0.099 but a $SENSI_{SPEC99\%}$ index of 29%. Finally, compound #05 provided noteworthy interesting indices, *i.e.* $AUC_{>70}$ 0.132 and $SENSI_{SPEC99\%}$ = 35%. Thanks to their *m/z* ratio and retention time values, these compounds could be hypothetically related to biologically pertinent steroid metabolites. Compound #04 could be associated with an isomer of AS, EtioS or DHTS. Moreover, previous results related to DHEA administration promote epiandrosterone sulphate as a good candidate [11]. Compound #05 could be putatively related to a hydroxylated form of AG, EtioG or DHTG. Mareck et al. reported that neither 11β-OH-androsterone glucuronide nor 11β-OH-etiocholanolone glucuronide urinary excretion levels were modified after testosterone intake [4]. On the other hand, known testosterone metabolites such as 16α-OH-androsterone glucuronide, 16α-OH-etiocholanolone glucuronide, 6β-OH-androsterone glucuronide or 6β-OH-etiocholanolone glucuronide constitute relevant possibilities [41]. Compound #06 corresponded to an isomer of either AG, EtioG or DHTG. Further investigations of compound #6 were undertaken with respect to these indications through comparisons with commercially available reference standards. Thanks to identical retention time and *m/z* values, this analyte was identified as DHTG. Since DHTG is a known parameter that is often part of the *steroid profile*, and was described as a sensitive and selective indicator of exogenous testosterone intake [42,43], its untargeted detection provides a noteworthy argument for the relevance of the presented approach.

These results suggested that both sulphate conjugates and unknown compounds # 04 and 05 may provide complementary reliable information in the perspective of multiparametric indicators. Indeed, higher detection abilities are expected through a longitudinal follow-up of these new parameters when compared with the present population-based criteria.

4. Concluding remarks

This study highlighted the opportunity to provide a deeper insight into the metabolic variations in urine after oral testosterone undecanoate intake thanks to an untargeted steroidomic approach. Data from a clinical study were assessed with N-PLS-DA, O-PLS-DA classification models and SUS-plot analysis. Urinary excretion kinetics were evaluated and used to distinguish between biomarkers related to either intermediate or marked phases of testosterone excretion after oral intake. This approach allowed relevant steroid metabolites candidates to be highlighted and confirmed with ROC curves inspection. Known metabolites of the *steroid profile* were analysed in a targeted manner and ensured the

relevance of this methodology by providing consistent results with the current biological knowledge. However, the validity of these results remains specific to the clinical study and a larger evaluation of the suitability of the candidates remains mandatory. This research has shown proof of principle but further studies are needed to validate the methodology before routine application. Moreover, a longitudinal follow-up of these biomarkers is expected to reveal their discriminating power with respect to reference sensitivity and specificity values applied in routine anti-doping control.

While multiparametric indicators are expected to provide evidence of AAS misuse with more confidence and handle effectively the intrinsic variability of biological origin and confounding factors, the discovery of potent parameters by untargeted strategies remains of utmost importance. The approach presented in this work constitutes therefore an attractive opportunity for the detection of biomarkers related to the administration of other AAS such as DHT or DHEA.

References

[1] P.E. Sottas, N. Robinson, M. Saugy, The athlete's biological passport and indirect markers of blood doping, in: D. Thieme, P. Hemmersbach (Eds.), Doping in Sports: Biochemical Principles, Effects and Analysis, vol. 165, Springer, Berlin, 2010, pp. 305–326.

[2] L. Malcovati, C. Pascutto, M. Cazzola, Hematologic passport for athletes competing in endurance sports: a feasibility study, Haematologica 88 (2003) 570–581.

[3] P.E. Sottas, M. Saugy, C. Saudan, Endogenous steroid profiling in the athlete biological passport, Endocrinol. Metab. Clin. 39 (2010) 59–73.

[4] U. Mareck, H. Geyer, G. Opfermann, M. Thevis, W. Schanzer, Factors influencing the steroid profile in doping control analysis, J. Mass Spectrom. 43 (2008) 877–891.

[5] M.K. Parr, W. Schanzer, Detection of the misuse of steroids in doping control, J. Steroid Biochem. 121 (2010) 528–537 (and references therein).

[6] World Anti-doping Agency (WADA), The World Anti-Doping Code. The 2011 Prohibited List, 2011 (accessed March 2011) www.wada-ama.org.

[7] M. Becchi, R. Aguilera, Y. Farizon, M.M. Flament, H. Casabianca, P. James, Gas-chromatography combustion isotope ratio mass-spectrometry analysis of urinary steroids to detect misuse of testosterone in sport, Rapid Commun. Mass Spectrom. 8 (1994) 304–308.

[8] E. Strahm, P.E. Sottas, C. Schweizer, M. Saugy, J. Dvorak, C. Saudan, Steroid profiles of professional soccer players: an international comparative study, Br. J. Sports Med. 43 (2009) 1126–1130.

[9] J.J. Schulze, M. Lorentzon, C. Ohlsson, J. Lundmark, H.K. Roh, A. Rane, L. Ekstrom, Genetic aspects of epitestosterone formation and androgen disposition: influence of polymorphisms in CYP17 and UGT2B enzymes, Pharmacogenet. Genomics 18 (2008) 477–485.

[10] J. Jakobsson, L. Ekstrom, N. Inotsume, M. Garle, M. Lorentzon, C. Ohlsson, H.K. Roh, K. Carlstrom, A. Rane, Large differences in testosterone excretion in Korean and Swedish men are strongly associated with a UDP-glucuronosyl transferase 2B17 polymorphism, J. Clin. Endocr. Metab. 91 (2006) 687–693.

[11] L. Dehennin, M. Ferry, P. Lafarge, G. Peres, J.P. Lafarge, Oral administration of dehydroepiandrosterone to healthy men: alteration of the urinary androgen profile and consequences for the detection of abuse in sport by gas chromatography mass spectrometry, Steroids 63 (1998) 80–87.

[12] L. Dehennin, P. Lafarge, P. Dailly, D. Bailloux, J.P. Lafarge, Combined profile of androgen glucuro- and sulfoconjugates in postcompetition urine of sportsmen: a simple screening procedure using gas chromatography mass spectrometry, J. Chromatogr. B 687 (1996) 85–91.

[13] C. Ayotte, Detecting the administration of endogenous anabolic androgenic steroids, in: D. Thieme, P. Hemmersbach (Eds.), Doping in Sports: Biochemical Principles, Effects and Analysis, vol. 165, Springer, Berlin, 2010, pp. 77–98.

[14] O.J. Pozo, P. Van Eenoo, K. Deventer, F.T. Delbeke, Detection and characterization of anabolic steroids in doping analysis by LC–MS, Trac-Trends Anal. Chem. 27 (2008) 657–671.

[15] E. Strahm, I. Kohler, S. Rudaz, S. Martel, P.A. Carrupt, J.L. Veuthey, M. Saugy, C. Saudan, Isolation and quantification by high-performance liquid chromatography-ion-trap mass spectrometry of androgen sulfoconjugates in human urine, J. Chromatogr. A 1196 (2008) 153–160.

[16] F. Badoud, E. Grata, J. Boccard, S. Rudaz, J.-L. Veuthey, M. Saugy, Quantification of glucuronidated and sulfated steroids in human urine by ultra-high pressure liquid chromatography quadrupole time-of-flight mass spectrometry, Anal. Bioanal. Chem. 400 (2011) 503–516.

[17] J. Boccard, J.L. Veuthey, S. Rudaz, Knowledge discovery in metabolomics: an overview of MS data handling, J. Sep. Sci. 33 (2010) 290–304.

[18] S. Anizan, D. Di Nardo, E. Bichon, F. Monteau, N. Cesbron, J.P. Antignac, B. Le Bizec, Targeted phase II metabolites profiling as new screening strategy to investigate natural steroid abuse in animal breeding, Anal. Chim. Acta 700 (2011) 105–113.

[19] I. Riedmaier, C. Becker, M.W. Pfaffl, H.H.D. Meyer, The use of omic technologies for biomarker development to trace functions of anabolic agents, J. Chromatogr. A 1216 (2009) 8192–8199.

[20] J.C.W. Rijk, A. Lommen, M.L. Essers, M.J. Groot, J.M. Van Hende, T.G. Doeswijk, M.W.F. Nielen, Metabolomics approach to anabolic steroid urine profiling of bovines treated with prohormones, Anal. Chem. 81 (2009) 6879–6888.

[21] N. Baume, Y.O. Schumacher, P.E. Sottas, C. Bagutti, M. Cauderay, P. Mangin, M. Saugy, Effect of multiple oral doses of androgenic anabolic steroids on endurance performance and serum indices of physical stress in healthy male subjects, Eur. J. Appl. Physiol. 98 (2006) 329–340.

[22] N. Baume, C. Saudan, A. Desmarchelier, E. Strahm, P.E. Sottas, C. Bagutti, M. Cauderay, Y.O. Schumacher, P. Mangin, M. Saugy, Use of isotope ratio mass spectrometry to detect doping with oral testosterone undecanoate: inter-individual variability of C-13/C-12 ratio, Steroids 71 (2006) 364–370.

[23] E. Strahm, Etude du métabolisme des androgènes et implications dans la lutte contre le dopage, Thèse n° 4185, Université de Genève, 2010.

[24] P.A. Guy, I. Tavazzi, S.J. Bruce, Z. Ramadan, S. Kochhar, Global metabolic profiling analysis on human urine by UPLC-TOFMS: issues and method validation in nutritional metabolomics, J. Chromatogr. B 871 (2008) 253–260.

[25] E. Fahy, S. Subramaniam, H.A. Brown, C.K. Glass, A.H. Merrill, R.C. Murphy, C.R.H. Raetz, D.W. Russell, Y. Seyama, W. Shaw, T. Shimizu, F. Spener, G. van Meer, M.S. VanNieuwenhze, S.H. White, J.L. Witztum, E.A. Dennis, A comprehensive classification system for lipids, J. Lipid Res. 46 (2005) 839–861.

[26] E. Fahy, S. Subramaniam, R.C. Murphy, M. Nishijima, C.R.H. Raetz, T. Shimizu, F. Spener, G. van Meer, M.J.O. Wakelam, E.A. Dennis, Update of the LIPID MAPS comprehensive classification system for lipids, J. Lipid Res. 50 (2009) S9–S14.

[27] C.M. Rubingh, S. Bijlsma, R.H. Jellema, K.M. Overkamp, M.J. van der Werf, A.K. Smilde, Analyzing longitudinal microbial metabolomics data, J. Proteome Res. 8 (2009) 4319–4327.

[28] R.A. Harshman, M.E. Lundy, Parafac – parallel factor-analysis, Comput. Stat. Data Anal. 18 (1994) 39–72.

[29] P. Geladi, Analysis of multi-way (multi-mode) data, Chemometr. Intell. Lab. 7 (1989) 11–30.

[30] R. Bro, Multiway calibration. Multilinear PLS, J. Chemometr. 10 (1996) 47–61.

[31] C.A. Andersson, R. Bro, The N-way Toolbox for MATLAB, Chemometr. Intell. Lab. 52 (2000) 1–4.

[32] L. Eriksson, J. Trygg, S. Wold, CV-ANOVA for significance testing of PLS and OPLS models, J. Chemometr. 22 (2008) 594–600.

[33] H.R. Norli, K. Esbensen, F. Westad, K.I. Birkeland, P. Hemmersbach, Chemometric evaluation of urinary steroid profiles in doping control, J. Steroid Biochem. 54 (1995) 83–88.

[34] L. Dehennin, A.M. Matsumoto, Long-term administration of testosterone enanthate to normal men – alterations of the urinary profile of androgen metabolites potentially useful for detection of testosterone misuse in sport, J. Steroid Biochem. Mol. Biol. 44 (1993) 179–189.

[35] A.T. Cawley, E.R. Hine, G.J. Trout, A.V. George, R. Kazlauskas, Searching for new markers of endogenous steroid administration in athletes: looking outside the metabolic box, Forensic Sci. Int. 143 (2004) 103–114.

[36] 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.

[37] J. Trygg, S. Wold, Orthogonal projections to latent structures (O-PLS), J. Chemometr. 16 (2002) 119–128.

[38] M. Bylesjö, M. Rantalainen, O. Cloarec, J.K. Nicholson, E. Holmes, J. Trygg, OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification, J. Chemometr. 20 (2006) 341–351.

[39] S. Wiklund, E. Johansson, L. Sjostrom, E.J. Mellerowicz, U. Edlund, J.P. Shockcor, J. Gottfries, T. Moritz, J. Trygg, Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models, Anal. Chem. 80 (2008) 115–122.

[40] F. Buiarelli, F. Coccioli, M. Merolle, B. Neri, A. Terracciano, Development of a liquid chromatography–tandem mass spectrometry method for the identification of natural androgen steroids and their conjugates in urine samples, Anal. Chim. Acta 526 (2004) 113–120.

[41] P. Van Renterghem, P. Van Eenoo, H. Geyer, W. Schanzer, F.T. Delbeke, Reference ranges for urinary concentrations and ratios of endogenous steroids, which can be used as markers for steroid misuse, in a Caucasian population of athletes, Steroids 75 (2010) 154–163.

[42] A.T. Kicman, S.B. Coutts, C.J. Walker, D.A. Cowan, Proposed confirmatory procedure for detecting 5-alpha-dihydrotestosterone doping in male-athletes, Clin. Chem. 41 (1995) 1617–1627.

[43] M. Donike, M. Ueki, Y. Kuroda, H. Geyer, E. Nolteernsting, S. Rauth, W. Schanzer, U. Schindler, E. Völker, M. Fujisaki, Detection of dihydrotestosterone (DHT) doping: alterations in the steroid profile and reference ranges for DHT and its 5 alpha-metabolites, J. Sports Med. Phys. Fitness (2011) 235–250.