Food Control 147 (2023) 109611

Contents lists available at ScienceDirect

Food Control

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

Potential and limitation of retrospective HRMS based data analysis: "Have meat-producing animals been exposed to illegal growth promotors such as SARMs?"

A. Kaufmann [a,1], A. Arrizabalaga-Larrañaga [b,1], M.H. Blokland [b,*], S.S. Sterk [b]

[a] Official Food Control Authority of the Canton of Zürich, Fehrenstrasse 15, 8032, Zürich, Switzerland
[b] Department of Growth Promotors, European Union Reference Laboratory, Wageningen Food Safety Research (WFSR), Part of Wageningen University & Research, 700 AE, Wageningen, the Netherlands

ARTICLE INFO

Keywords:
Selective androgen receptor modulators
Growth promotors
High-resolution mass spectrometry
Retrospective analysis
Cattle

ABSTRACT

Targeted strategies using gas or liquid chromatography coupled to triple quadrupole mass analyzers are mostly used in food safety control laboratories. However, when a new chemical hazard emerges, the target list needs to be updated and thus the should be reanalyzed. Under this scenario, the constant emergence of compounds that can illegally be used as growth promotors in cattle, such as selective androgen receptor modulators (SARM), and the phasing out of older SARMs, would require routine control laboratories (RCL) to constantly re-develop and re-validate their once developed targeted methods. Nevertheless, high-resolution mass spectrometry (HRMS) is not limited to analyzing samples using only a predefined targeted list and therefore the implementation of untargeted HRMS methods in RCLs to detect SARMs would enhance the laboratories throughput. In this study, two HRMS profile databases of cattle urine samples gathered over the past years and collected by two independent laboratories, were retrospectively analyzed to find out if a new SARM has been used. This study assesses different retrospective screening approaches based on the Schymanski identification confidence levels. A screening purely based on identification confidence levels 4 and 5 does not permit reliable detection of any exogenous trace compound that does not belong to a certain food or environmental matrix sample due to the high number of false detects. Therefore, it is shown that the availability of a physical reference substance to increase the identification confidence up to level 1 is essential to discriminate a suspected detection between a false positive detection or a confirmed finding.

1. Introduction

Most of the current analytical methods used for food safety research are targeted to a list of compounds employing gas or liquid chromatography coupled to triple quadrupole mass analyzers (Steiner, Malachová, Sulyok, & Krska, 2021). In case a new chemical hazard emerges, in control monitoring plans, this compound must be added to the target list of the used method, and therefore the sample need to be reanalyzed. However, reanalyzing samples is time-consuming and often impossible due to limited storage facilities in control laboratories. Furthermore, it is unknown if analytes of interest remain detectable after a long cold storage or after repeated freezing thawing cycles. High-resolution mass spectrometry (HRMS) is not limited to analyze (food) samples using only a predefined targeted list (Yan, Zhang, Zhou, Li, & Feng, 2022). HRMS measurements can collect full-scan mass spectra with a high mass resolution, typical >20,000 FWHM, and a mass accuracy better than 3 ppm. The collected data can be processed in a targeted way using a list of m/z and retention times of known compounds or retrospectively to search new compounds that were not included in the original targeted list of compounds (Jongedijk et al., 2023). In the last decade, numerous papers have described the possibility of retrospectively searching HRMS data files for veterinary drug (Jongedijk et al., 2023; Xu, Li, Chen, Ma, & Yang, 2018), pesticides, antibiotics and mycotoxins (De Dominicis, Commissati, Gritti, Catellani, & Suman, 2015; Gómez-Pérez, Romero-González, Vidal, & Frenich, 2015; Polgár et al., 2012) identification. In this way, if the HRMS data is acquired correctly, data files that were

* Corresponding author.
E-mail address: marco.blokland@wur.nl (M.H. Blokland).
[1] authors contributed equally.

https://doi.org/10.1016/j.foodcont.2023.109611
Received 26 August 2022; Received in revised form 3 January 2023; Accepted 5 January 2023
Available online 9 January 2023

collected years before can still be used for retrospective identification of new emerged hazards. Therefore, HRMS-based retrospective screening approach could be useful in food safety research to detect, for example, the potential misuse of emerging growth promoters such as selective androgen receptor modulators (SARMs) in animal husbandry.

SARMs are anabolic non-steroidal compounds well known to be used as growth promoters in both human- and animal sports due to their easy availability through the black market (Kohler et al., 2010; Van Wagoner, Eichner, Bhasin, Deuster, & Eichner, 2017), simplicity of use (oral administration though pills or powders), and advantageous biological effects (Segal, Narayanan, & Dalton, 2006). These facts could lead to their misuse in cattle to increase muscle mass and farmers profits. For this reason, SARMs are classified as banned compounds in live animals and products of animal origin in EU Council Directive 96/22/EC (2008) and Regulation (EU) 2022/1644 (2022). The concern in this field has led to the development of several screening (Arrizabalaga-Larrañaga, Nielen, & Blokland, 2021; Geldof et al., 2017; Temerdashev et al., 2017) and confirmation (Gadaj et al., 2020; Gadaj, Ventura, Ripoche, & Mooney, 2019; Schmidt & Mankertz, 2018; Ventura et al., 2019; Ventura, Gadaj, Buckley, & Mooney, 2020) analytical methods for the analysis of SARMs in many matrices. Most of these methods are based on liquid chromatography coupled to tandem mass spectrometry (LC−MS/MS) and are used for the analysis of dietary supplements and animal urine samples. Nevertheless, today, the detection of these compounds in biological samples such as blood and urine continues to be a problem of great importance for residue control laboratories (RCL), since it is known that SARMs undergo metabolism to phase I metabolites, but also phase II conjugates are generated (de Rijke et al., 2013; Hansson et al., 2015; Thevis et al., 2010). Data on the metabolism of SARMs in bovine animals are scarce. In a study performed by Rojas et al. bicalutamide was administered to a calf, urine and feaces were collected, bicalutamide was found to be the major excreted compound in its free form in urine and feaces (Rojas et al., 2017). The concentrations in feces were higher, but the overall detection time window was similar, so both matrices can be used for residue control. The observed metabolites were either hydroxylation (phase I reaction) or glucuronidation and sulfation (phase II reactions). For enobosarms the excretion profile for urine and feaces was different (Cesbron et al., 2017). In both matrices, the administrated compound was the major compound present and unconjugated. However, in feaces the detection window was 21 days compared to 9 days in urine. In another study by de Rijke et al. similar results were reported for the extrected metabolites of ostarine in urine, whereby the administrated compound was the most abundant compound present in urine (de Rijke et al., 2013). Based on these studies, the unconjugated compound is found to be a suitable marker for administration in bovine. Urine and feaces can be used as a matrix for detection, whereby feaces have a longer detection window in one study. However, urine in another study has a similar detection window, so urine can be considered a suitable matrix to detect SARMs.

Regular new SARM are developed in this field. Unfortunately, right after developing these new SARMs, they become quickly available on the black market, where they can be misused in sports doping or in animal husbandry. The constant emergence of new SARMs, and the phasing out of older SARMs, would require RCLs to constantly re-develop and re-validate their once developed control methods. For this reason, it is interesting to implement untargeted HRMS methods in RCLs to detect SARMs. Within this HRMS workflow, the data analysis could be easily adjusted to detect new SARMs, and in case a database of HRMS profiles is available, samples gathered over the past years could be analyzed by retrospective analysis to find out if a new SARM is identified. Therefore, the present study aimed to demonstrate and discuss several approaches by employing different identification confidence levels of retrospective data analysis from HRMS data of residue control programs using SARMs as model compounds.

2. Materials and methods

2.1. Reagents and standards

2.1.1. Wageningen Food Safety Research (WFSR)

Ethanol absolute (EtOH), *tert*-Butyl methyl ether (TBME), disodium phosphate dihydrate (Na_2HPO_4 $2H_2O$), sodium hydroxide (NaOH) all for analysis EMSURE® quality degree were purchased from Merck KGaA (Darmstadt, Germany), Methanol and acetonitrile HPLC Supra-gradient quality were bought from Biosolve B.V. (Valkenswaard, The Netherlands) and J.T. Baker (Reinach, Switzerland). β-Glucuronidase E Coli K12 was purchased from Roche CustomBiotech (Basel, Switzerland).

Analytical standards of three SARMs including, ostarine, andarine and bicalutamide, were purchased from Selleck chemicals (Houston, Texas, United States), whereas bicalutamide-d4 from Toronto Research Chemicals (Toronto, Canada). Individual stock solutions (1,000 mg L^{-1}) of ostarine, andarine and bicalutamide were prepared in ethanol and stored at −80 °C. Intermediate mixture standard solutions (10, 1, 0.1, and 0.01 mg L^{-1}) containing all target compounds were prepared from stock standard solution and their consecutive dilution by appropriate dilution in ethanol. Working standard solutions (1 μg L^{-1}) were prepared from appropriate dilution in water:acetonitrile (80:20, *v/v*). All these standard solutions were stored at −80 °C until their use.

2.1.2. Official Food Control Authority of the Canton of Zürich (KLZH)

Dimethylsulfoxide, ethyl acetate and potassium carbonate were obtained from VWR (Dietikon, Switzerland), acetic acid glacial from Scharlau (Barcelona, Spain) and ammonium fluoride and sodium acetate from Sigma-Aldrich (Busch, Switzerland). BMS 564929 (2-Chlor-3-methyl-4-[(7R,7aS)-tetrahydro-7-hydroxy-1,3-dioxo-1H-pyrrolo[1,2-c] imidazole-2(3H)-yl]benzonitrile) was purchased from Sigma-Aldrich (Buchs, Switzerland) whereas GLPG0492 (4-[(4S)-4-(Hydroxymethyl)-3-methyl-2,5-dioxo-4-phenyl-1-imidazolidinyl]-2-(trifluoromethyl)benzonitrile), LDG-4033 (4-[(2R)-2-[(1R)-2,2,2-trifluoro-1-hydroxyethyl]-1-pyrrolidinyl]-2-(trifluoromethyl)-benzonitrile), S-1 (4-Desacetamido-4-fluoro andarine) and S-9 (4-Desacetamido-4-chloro andarine) were purchased from Chemie Brunschwig (Basel, Switzerland).

Individual stock standard solutions of BMS 564929, GLPG0492, LDG-4033, S-1 and S-9 were prepared in acetonitrile at 1,000 mg $L^{-1,}$ and the mixed spiking solution (0.5 mg L^{-1}) was obtained from the dilution of stock standard solutions in acetonitrile and stored at −15 °C until its use.

2.2. Samples

Samples used for the ring study organized by the European Union Reference Laboratory for residues were prepared by spiking three urine samples with ostarine, andarine and bicalutamide at 5 ng mL^{-1}.

1457 samples were used for retrospective LC−HRMS analysis by the Official Food Control Authority of the Canton of Zürich (KLZH). Among them, 35% were urine (bovine and pig), 34% whole blood (bovine), and 31% liver (bovine). All these samples originated from Switzerland. They were collected by trained veterinarians, either from the slaughtering facilities (liver and urine) or directly from the farm (blood). The sampling time period was extended over three years. All samples were immediately frozen after the completion of the sampling and analyzed within a period of less than one month.

1300 urine samples were used for retrospective LC−HRMS analysis by the European Union Reference Laboratory for residues at Wageningen Food Safety Research (WFSR). The urine samples used were collected over eight years (2014–2021) on farms in The Netherlands. These urines used for this study cannot be traced back to their origin.

2.3. Sample analysis and LC-MS conditions

Sample sets were analyzed following different methods (sample treatment and LC-MS systems) as described in the supplementary material (Fig. S1). Briefly, the spiked samples prepared by EURL Wageningen were analyzed by Method 1, the 1457 samples analyzed by KLZH for retrospective analysis followed Method 2 and the 1300 urine samples analyzed by WFSR used Method 3.

3. Results and discussion

3.1. Evaluating the prerequisites for retrospective analysis

The performance of the HRMS workflow was assessed by using samples from a ring study organized by the European Union Reference Laboratory (EURL) for residues located in Wageningen (The Netherlands) on the analysis of SARMs in urine. The EURL invited all European National Reference Laboratories (NRL) responsible for analyzing growth promotors in animal based food to participate in a quantification and confirmation study. To check the spikes amount of the ring test materials Method 1 (described in supplementary material, Table S1) was used. This method has been validated by EURL-Wageningen, showing good linearity of the analytical response of target compounds within the working range with correlation coefficients (r) higher than 0.990. Additionally, the method demonstrates to be selective and robust, with high accuracy (≥99%) and good repeatability, as it is summarized in Table S2. The spiked urine ring test material included three SARMs; ostarine, andarine and bicalutamide.

The prepared ring test samples were sent to ten NRLs to analyze them with their in-house methods. After analysis, the consensus value for the target compounds was 4.5 ng mL^{-1} for Ostarine, 4.7 ng mL^{-1} for Andarine and 4.6 ng mL^{-1} for Bicalutamide. As can be shown, all of the participating laboratories produced results that were in good agreement with the values assigned by the EURL Wageningen (Fig. S2).

Based on these initial results, the NRL from the Official food control authority of the canton of Zürich realized that the used method for the ring test showed high similarities with a multi-residue method (Method 2) routinely used in the KLZH to analyze growth promotors. In this case, the last method covers 43 analytes belonging to natural and synthetic steroids, zeranols and stilbenes. Besides, this method has also been validated for a number of matrices such as liver, muscle, urine, blood and serum as described by Kaufmann (Kaufmann, 2020). In this way, both methods use an enzymatical cleavage step and a similar liquid/-liquid extraction procedure, but mass analyzer used is different, being the first method carried out by a low-resolution triple quadrupole in multiple reaction monitoring mode (MRM) and the second one by a high-resolution Quadrupole Orbitrap in full scan mode. The high similarity between the two methods encourages KLZH to evaluate the capability of Method 2 to analyze SARMs. Hence, for this purpose, blank urine samples were fortified with the provided SARMs standard mixture and analyzed by employing Method 2. The ad-hoc validation showed that Method 2 is capable of not only detecting but also quantifying the three compounds of interest (andarine, ostarine and bicalutamide). This observation finally led to the idea to reprocess stored data files linked to routine samples previously analyzed for the presence of growth promotors in KLZH with Method 2. Successfully using a particular analytical method for retrospective purposes certainly requires a number of intrinsic method performance qualities. The method is expected to be sensitive, and that a linear signal abundance (peak area) versus analyte concentration relationship exists. In addition, the method has to show long-term stability. This, unfortunately, can not be proven for the analytes of retrospective interest. Yet the method has been validated for a set of related compounds (illegal growth promotors) at a similar concentration range for the matrices of interest. The validation data (Kaufmann, 2020) lists the performance criteria for some 40 analytes. In addition, the whole validation series (based on different, repeated fortification levels) was repeated three times (different sample processing days) by different analytes to mimic the variation occurring within a historical time range.

Within WFSR, at the same time, there was also a program in place whereby HRMS urine profiles were acquired using a Q-Orbitrap mass analyzer. In this case, approximately every year, 250 samples are analyzed by HRMS with Method 3. As the sample clean-up of the method is generic, it was expected that the method was capable of detecting SARMs as well. Therefore, first, samples were spiked with ostarine, andarine and bicalutamide and were analyzed. The three compounds could be easily detected in these spiked samples, demonstrating that the method can detect these SARMs in the samples. Since these three SARMs can be detected in the spiked samples, which were similarly processed as all other samples in the archive of bovine urine HRMS profiles collected over the past years, this archive can be retrospectively searched for at least these three SARMs and also for similar structural SARMs. It must be noted that for method 3 no deconjugation is performed on the urine samples. The lack of deconjugation enables the measurement of phase II metabolites directly with method 3. Studies conducted on the metabolism of SARMs in bovine demonstrated that the major compound excreted is the administrated compound in its free form. When new information on the metabolism of SARMs in bovines becomes available, the archive can be searched again for these metabolites.

3.2. Retrospective analysis aided by the availability of physical reference substances

The selectivity provided by the accurate mass of HRMS, the narrow mass extraction window and the knowledge of the analyte retention were sufficient to screen for the three analytes (ostarine, andarine and bicalutamide) not only in urine, but also blood and liver samples. Hence, the data files of all samples analyzed with Method 2 and 3 were retrospective analyzed. Fig. 1 shows the extracted LC−HRMS chromatograms of (i) ostarine, (ii) andarine and (iii) bicalutamide at a concentration of 2 μg L^{-1} by Method 2. The top row (Fig. 1A) shows the signals obtained when injecting the standard solution while the second row (Fig. 1B) shows a typical urine sample processed by Method 2 where no relevant signal was observed. The third row (Fig. 1C) shows selected urine samples with produced “false detected” signals and the bottom row (Fig. 1D) depicts the signal obtained when analyzing a liver sample. Note, in all cases, the y-axis was normalized (the basis for the normalization was the peak height produced by using the 2 μgL^{-1} standard solutions) for each compound to permit an optical (signal intensity) comparison. The term “false detected”, as used here, refers to the fact that a relevant chromatographic peak was noted, but the retention time of that peak was significantly diverging from the expected analyte retention time and thus, considering it as our interest compound, it would deliver to false positive detection. It is important to notice that most analyzed urine samples produced no discernible signal at all, as shown in Fig. 1B. The second type of urine sample (Fig. 1C) reflects the most intensive false positives encountered when doing retrospective analysis of more than 1000 historical samples. Liver behaves significantly different from urine. Liver samples do not only produce more frequent false positives than urine samples, but liver false positives produce clearly higher ion abundances (peak areas). Although the signals present in the liver samples were more frequent and abundant, all could be discarded as false detected based on the retention time.

For Method 3, similar results were obtained as with Method 2, the combination of accurate mass, narrow mass extraction windows, and the comparison of retention times eliminated any possible false-positive results. From the obtained results, none of the investigated samples with Methods 2 or 3 showed the presence or suspected presence of ostarine, andarine and bicalutamide.

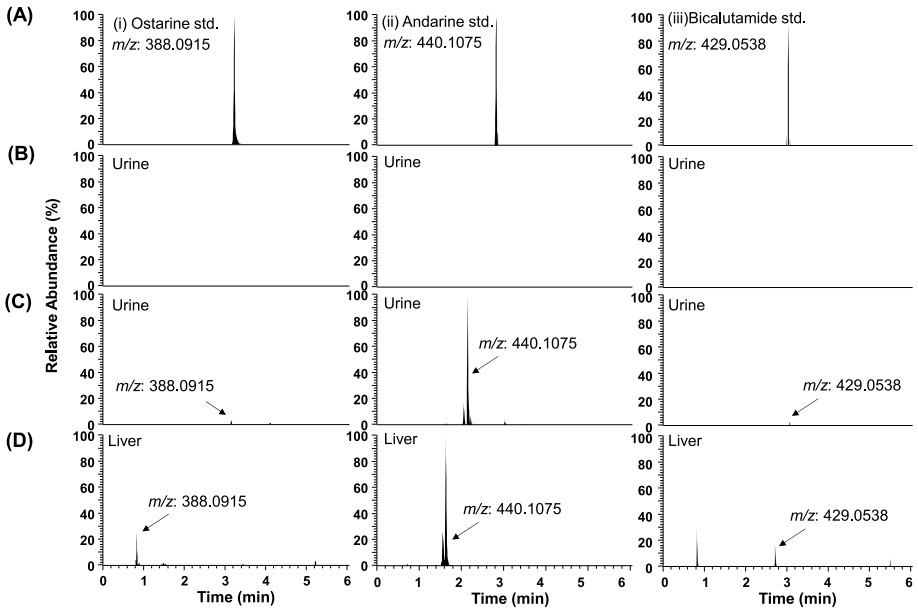

Fig. 1. LC−HRMS chromatogram obtained from (A) standard solution of (i) ostarine, (ii) andarine and (iii) bicalutamide at a concentration of 2 μg L^{-1}; (B) urine sample, (C) urine samples with produced false positives and (D) liver sample employing method 2.

3.3. Retrospective analysis

Although the targeted three SARMs were not detected, all data files of Method 2 were retrospectively analyzed for the presence of signals belonging to another set of potentially used SARMs. See Fig. 2 for the chemical structures of SARMs used in this retrospective search. To classify the retrospective results, we used the confidence levels proposed by Schymanski et al. (Schymanski et al., 2014). At the time of retrospective analysis by LC−HRMS, it must be pointed out that no physical reference substances were available. Therefore, according to Schymanski et al. (Schymanski et al., 2014) only a maximum identification confidence level 2 can be obtained.

3.3.1. Retrospective analysis based on the use of exact mass only

The retrospective search was first performed using only the exact mass (identification confidence level 5). A significant number of only marginally mass deviating (less than 5 ppm) peaks remained Fig. 3 shows examples of four suspected SARMs (columns). The second row (B) gives a typical urine sample, showing no signals if retrospective search is performed for the exact mass and can be declared as negative for the searched compounds. The third row (C) shows urine samples where significant suspected analyte signals were encountered. As observed for the three previously targeted SARMs, the liver samples (forth row, D) showed more and most intense signals when only the exact mass was used. The depicted liver sample in the first columns even produced three relevant signals of LGD-4033(*m/z* 338.0854). As mentioned above, the analytical standards were not available at the time of the retrospective analysis, hence it would be impossible to reach a high identification confidence level.

Having no access to physical reference substances means that mass spectrometry-based detection has to rely on a number of assumptions. This begins with the selection of the polarity of ionization. In the case of

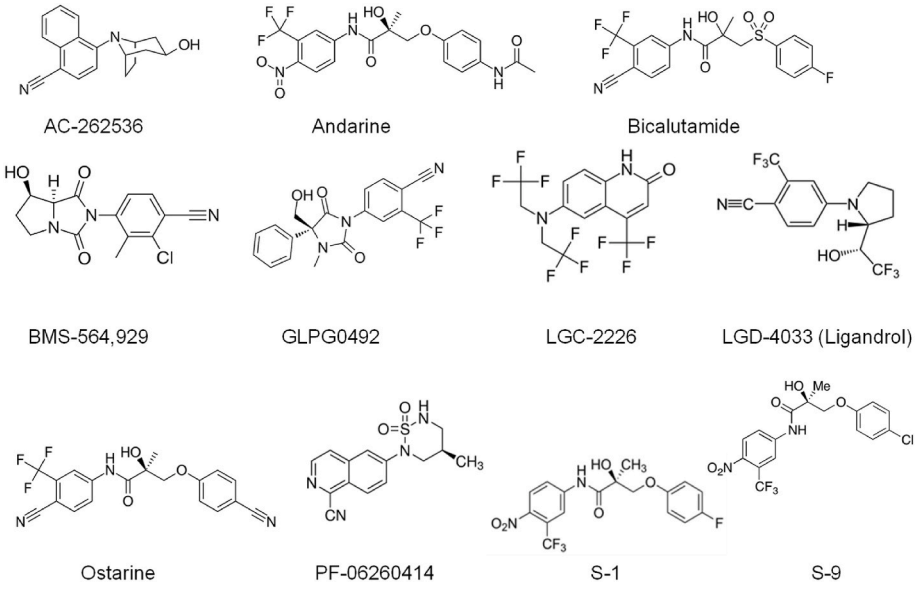

Fig. 2. Chemical structure of targeted and suspected SARMs.

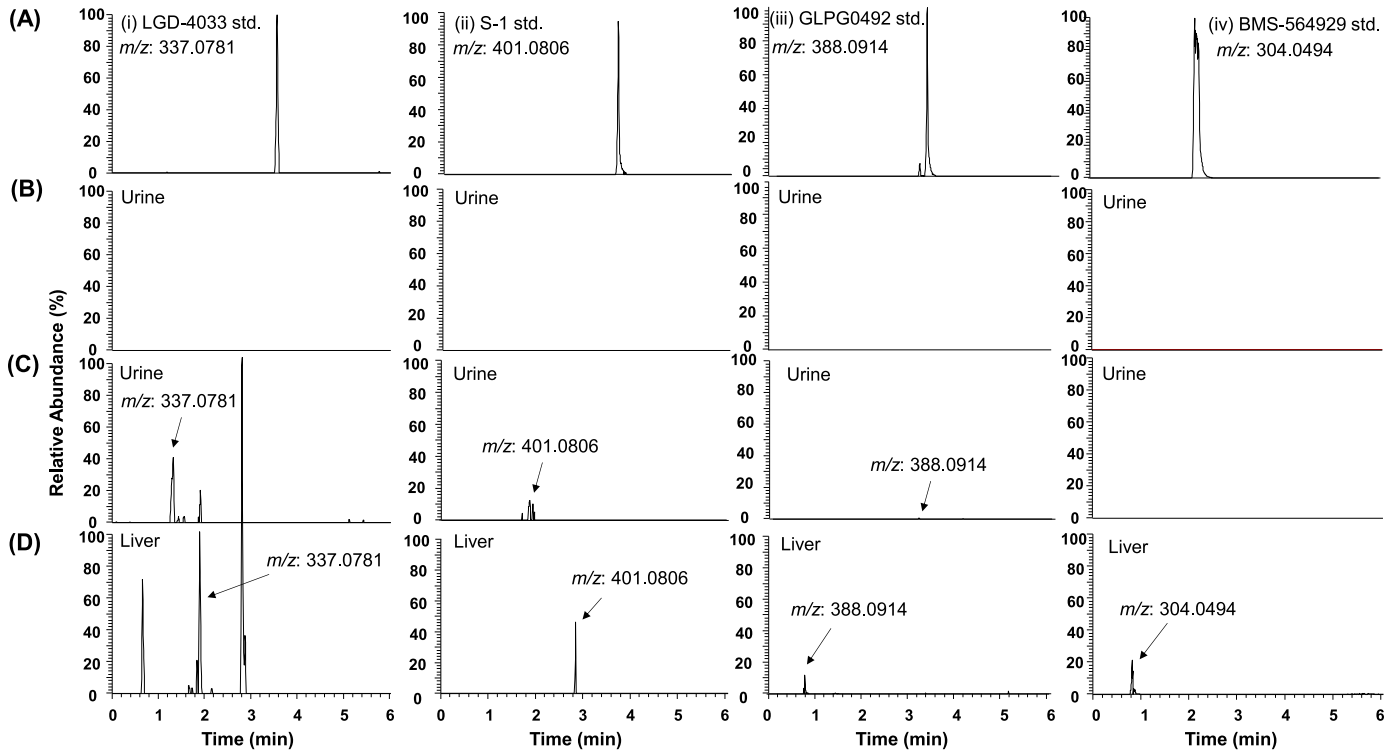

Fig. 3. Extracted ion chromatograms of the exact masses of (i) LGD-4033, (ii) S-1, (iii) GLPG0492 and (iv) BMS-564929 for (A) standard solution, (B) typical urine sample, (C) dirty urine samples, (D) liver sample employing method 2.

SARMs, at first point, it was assumed that all compounds produce negative ions since for the EURL-Wageningen method (Method 1) all analytes showed intense signals in the negative ionization mode. Therefore, the chemical structures of these potential SARMs were evaluated (Fig. 2), and they lead to the presumption that all these compounds should ionize when using negative ion mode by ESI. However, other research studies used negative as well as positive ionization for their analysis (Gadaj et al., 2020; Ventura et al., 2019). Although no systematic comparison is available, it has to be concluded that the obtained ion abundance (using positive *vs*. negative) ionization mode is not only determined by the chemical structure of the analyte, but also on the composition of the mobile phase, the interface parameters, and frequently also on the type (brand) of used ESI interface. Hence, the researcher does not know if the selected detection parameters enable sufficiently sensitive detection of each suspected analyte of interest. Therefore, this implies that not only intense but also small chromatographic peaks should be carefully investigated.

3.3.2. Retrospective analysis based on the use of exact mass and isotopic ratio

Besides the exact mass there are two other types of mass spectrometry based information available for retrospective analysis; the isotopic ratio and the structural information derived from product ions. Isotopic patterns can easily and accurately be calculated. This refers not only to the masses but also to the relative abundances. The use of isotope ratio will increase the identification confidence level to 4 (Schymanski et al., 2014). The relative isotopic ratios of the intact precursor ion were used to evaluate if a particular chromatographic peak can be ruled out to be a false positive level 5 identification. For instance, it appeared that in a number of cases the observed false positive is not a precursor ion but actually the first isotopic ion of a compound with a mass of $[M-H]^-$1.0033. An example of such false positive is shown in Fig. 4. The first isotope of a matrix peak has an exact mass which is similar to ostarine and this matrix peak occurs in both urine (Fig. 4C) and liver samples (Fig. 4D).

Isotopic signals of most analytes are significantly smaller than the corresponding analyte monoisotopic ion. This creates a major problem in the field of trace analysis. First, the ion abundance may be close to the physical detection capability of the instrument. More often, one or several interfering matrix ions may significantly modify the observed relative isotopic ratio. Encountering a detection where the ratio of the first isotopic peak is clearly below that of the expected spectra permits the conclusion of a false positive detection. Yet this is not true if a feature is encountered where the first isotopic peak abundance is higher than the theoretical spectra of the targeted analyte. Such an observation may be explained by the presence of the suspected analyte and the coelution of a matrix compound which is isobaric with the first isotopic peak of the suspected compound. Looking at the density of mass peaks clustered around every integer mass, it shows that isobaric interference at low ion abundances are not rare events. Furthermore, isotopic ratios of compounds containing only C; H; N and O atoms do not produce very unique isotopic patterns. The limited accuracy of absolute or relative ion abundance measurements in combination with co-eluting isobaric compounds limit the diagnostic value of isotopic ratios. An example of a false positive detection due to isobaric interference is shown in Fig. 4. In a liver sample (Fig. 4D), lots of interfering peaks are visible for S-1 causing a false positive signal. Therefore, the examples given in Fig. 4 demonstrate that even the combination of mass deviations and isotopic ratios was unable to filter out a false positive. Many SARMs contain atoms (F or Cl) that cause relevant mass defects. Using these mass effects for identification has shown to be a promising strategy, as demonstrated by (Léon et al., 2019) and could be used for those SARMs which contain one or more chlorine atoms. Chlorine does not only show a significant mass defect but, in addition, is responsible for a unique, easily detectable, isotopic pattern. This is much less the case for fluorine which occurs in nature without isotopes. Fluoride has a mass defect, yet even the presence of many fluorine atoms (as present in poly-fluoroalkyl substances) makes the mass defect alone a rather weak filtering tool (Kaufmann, Butcher, Maden, Walker, & Widmer, 2022).

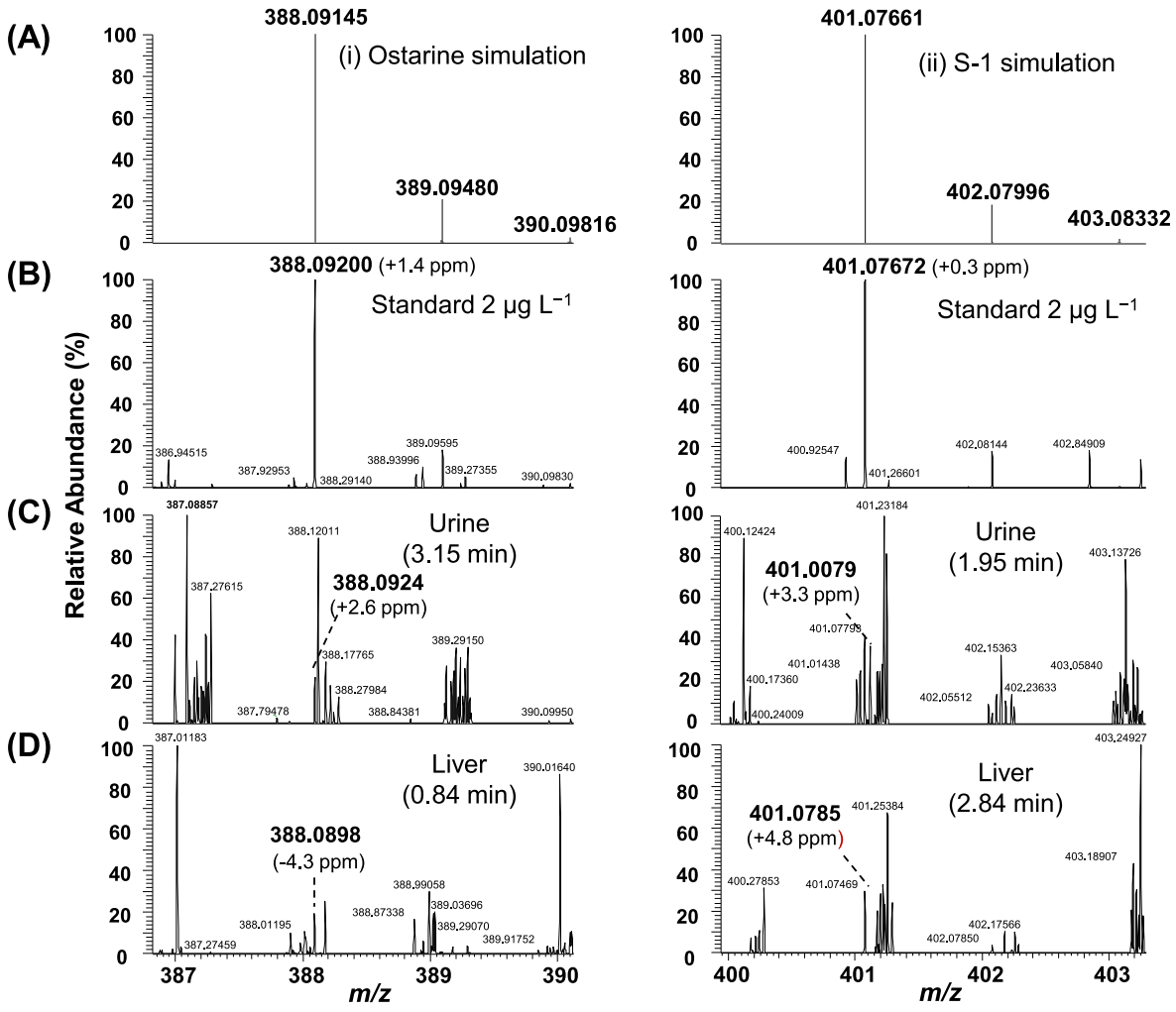

Fig. 4. HRMS spectra of (i) ostarine and (ii) S-1, (A) simulated, (B) a standard solution, and false positive detection in (C) urine sample, (D) liver sample.

3.3.3. Retrospective analysis based on the use of ion fragmentation information

The second mass spectrometry based way to address the issue of false positive detections is the use of product ion spectra and increase the identification confidence level to 3 or even 2 (Schymanski et al., 2014). Again, in the absence of physical reference substances, product ion spectra may be obtained from spectra libraries or proposed by in-silico fragmentation. Both strategies generally rely on the use of a unit mass isolated precursor which after isolation undergoes collision induced fragmentation. Such a dataset is generally not or limited available when attempting to perform retrospective analysis. At the time of analysis, neither the mass of the precursor ion of interest, not a suitable collision energy was known. A more feasible approach may be the use of a so called all ion fragmentation (AIF) trace (Zomer, Mol, & UR, 2020). Yet AIF of trace analytes in complex matrices may be of limited help in many cases. The use of a ramped collision energy may assure that each analyte of interest shows some fragmentation. Yet, this is also true for any co-eluting matrix compounds. In most cases low abundance analyte product ions are buried within a dense background of matrix compound fragments. Using in-silico fragmentation will propose likely product ions for the suspected analyte, yet the dozens of matrix related fragments crowded around each integer *m/z* will obscure these masses. A step forward is the deconvolution of a set of low and high energy mass traces and the alignment of these two traces. This technique commercialized by Waters was termed MSE. The alignment can be further improved by the inclusion of an ion mobility dimension (Kaufmann, Butcher, Maden, Walker, & Widmer, 2017). Unfortunately the dataset available for retrospective analysis was not acquired with an instrument having ion mobility.

3.3.4. Importance of physical reference substances (retention time)

At a later stage, the physical reference substances were available to increase the identification confidence level to level 1 and provide an unequivocal answer. Retention times can accurately be determined, but probably and more importantly, they are highly orthogonal to mass spectrometry derived information. After analyzing the reference standards, the previously obtained retrospective results based on mass spectrometry data only were re-assed. By adding the retention time obtained from the standards, many identified peaks could be discarded, as demonstrated in Fig. 3. The bottom row shows a chromatogram of a mixed standard solutions produced by injecting a diluted solution (2 µg L^{-1}) of the physical reference substances. It is immediately clear that the previously identified peaks are eluting at a different retention time and thus, they can be discarded. After retrospective analysis of all samples using reference standards, exact mass, and isotope distribution, no positive signals remained with an identification confidence level 1 retrospective search.

3.3.5. "Plausibility" a strategy to discard false detects

Attempting retrospective analysis requires a data set that was produced under identical conditions. Changing the sample preparation or the detection methodology within the historical measurements time series will complicate any retrospective work. Knowing more about a sample than just a sample number is important such as categorized

sample information (e.g. animal species, type of organ, sex, age of the animal, conventional or organic production, country of origin etc). This may permit the conclusion that all the samples belonging to a particular subset (e.g male bovines) show a particular chromatographic peak at a given retention time and that the concentration distribution rather resembles a normal distribution and not an outliner plot. Finding a particular signal in all samples within a sub-sample points to a false detection. This will be even more the case if the animals or plants were bred or grown in completely different geographical locations or growing regimes (e.g. conventional or organic production). On the other hand, encountering outliners within a particular sample sub-category should lead to investigating the sample with the highest concentration of the suspected compound. The high signal intensity permits a more reliable investigation of the monoisotopic mass, isotopic pattern, and mass defects. If the sample is still available, a reprocessing and production of a unit mass-isolated product ion spectra should be attempted.

4. Conclusions

Retrospective analysis of existing data sets may be an interesting tool when it comes to evaluating the need for extending the scope of a currently monitored set of analytes. Yet, integrating further compounds into an existing method does not only require the addition of them to the method, but may also lead to the need for a re-development of the method. Moreover, this fact would likely have to be followed by some sort of re-validation of the used method. Such issues become even more complex in the case of multi-national control plans because involved laboratories could use different equipment and may have a different level of training. Therefore, a prior evaluation of the likelihood that new analytes can be found in real samples and the required efforts to integrate them in existing methods is important. However, resources (either personal, instrumental or monetary) available for food safety monitoring are limited. This means attending to a new set of compounds may lead to the reduction of analytical activities in other fields of food safety monitoring. From that point of view, the application of the retrospective analysis may be an essential tool. A retrospective study on finding suspect compounds based on identification confidence levels 1 to 5 as reported in this paper should conclude that they have likely not been used or abused within the restricted set of retrospectively investigated samples. Hence, it is questionable if including these compounds in extensive monitoring programs makes sense. On the other hand, retrospective analysis can be a powerful tool when it comes to the early recognition of an emerging food scandal. Nevertheless, there is the likelihood of detecting such compounds by HRMS retrospective analysis, and hence, they may be detected before a large-scale food scandal is developed. Yet, this requires a limited number of laboratories that employ HRMS-based generic analytical methods in order to permit retrospective analysis. Finding such compounds in routine samples would justify the integration of these compounds in routine (e.g. triple quadrupole based) monitoring programs. As demonstrated in this study, a screening purely based on identification confidence levels 4 and 5 does not permit reliable detection of any exogenous trace compound that does not belong to a certain food or environmental matrix sample due to the high number of false detects. Suspecting a certain compound or a set of compounds based on level 2, 3, 4 and 5 may not yet lead to its's unambiguous detection. The availability of a physical reference substance to increase the identification confidence to level 1 is essential to discriminate a suspected detection between a false positive detection or a confirmed finding.

Fundings

This project was financially supported by the European Commission DG Health and Food Safety (EURL) and the Dutch Ministry of Agriculture, Nature and Food Quality under their statutory tasks program; WOT-02-003-066.

Compliance with ethical standards

Not applicable.

CRediT authorship contribution statement

A. Kaufmann: Conceptualization, Methodology, Investigation, Writing – original draft. **A. Arrizabalaga-Larrañaga:** Conceptualization, Methodology, Investigation, Writing – original draft. **M.H. Blokland:** Conceptualization, Methodology, Writing – review & editing. **S.S. Sterk:** Supervision.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors wish to acknowledge Bert Brouwer and Ingrid Elbers for the organization of the EURL training and the preparation of the samples.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodcont.2023.109611.

References

Arrizabalaga-Larrañaga, A., Nielen, M. W. F., & Blokland, M. H. (2021). Hand-held diode laser for on-site analysis using transportable mass spectrometry. *Analytical Chemistry, 93*(23), 8122–8127. https://doi.org/10.1021/acs.analchem.1c01083

Cesbron, N., Sydor, A., Penot, M., Prevost, S., Le Bizec, B., & Dervilly-Pinel, G. (2017). Analytical strategies to detect enobosarm administration in bovines. *Food Additives & Contaminants: Part A, 34*(4), 632–640. https://doi.org/10.1080/19440049.2016.1258122

Commission Delegated Regulation (EU) 2022/1644 of 7 July 2022 supplementing Regulation (EU) 2017/625 of the European Parliament and of the Council with specific requirements for the performance of official controls on the use of pharmacologically active substances authorised as veterinary medicinal products or as feed additives and of prohibited or unauthorised pharmacologically active substances and residues thereof (Text with EEA relevance).(2022).

Council Directive 96/22/EC of 29 April 1996 concerning the prohibition on the use in stockfarming of certain substances having a hormonal or thyrostatic action and of β-agonists, and repealing Directives 81/602/EEC, 88/146/EEC and 88/299. (2008). EEC.

De Dominicis, E., Commissati, I., Gritti, E., Catellani, D., & Suman, M. (2015). Quantitative targeted and retrospective data analysis of relevant pesticides, antibiotics and mycotoxins in bakery products by liquid chromatography-single-stage Orbitrap mass spectrometry. *Food Additives & Contaminants: Part A, 32*(10), 1617–1627. https://doi.org/10.1080/19440049.2015.1061703

Gadaj, A., Ventura, E., Healy, J., Botrè, F., Sterk, S. S., Buckley, T., et al. (2020). Enhanced UHPLC-MS/MS screening of selective androgen receptor modulators following urine hydrolysis. *MethodsX, 7*, Article 100926. https://doi.org/10.1016/j.mex.2020.100926

Gadaj, A., Ventura, E., Ripoche, A., & Mooney, M. H. (2019). Monitoring of selective androgen receptor modulators in bovine muscle tissue by ultra-high performance liquid chromatography-tandem mass spectrometry. *Food Chemistry X, 4*, Article 100056. https://doi.org/10.1016/j.fochx.2019.100056

Geldof, L., Pozo, O. J., Lootens, L., Morthier, W., Van Eenoo, P., & Deventer, K. (2017). In vitro metabolism study of a black market product containing SARM LGD-4033. *Drug Testing and Analysis, 9*(2), 168–178. https://doi.org/10.1002/dta.1930

Gómez-Pérez, M. L., Romero-González, R., Vidal, J. L. M., & Frenich, A. G. (2015). Identification of transformation products of pesticides and veterinary drugs in food and related matrices: Use of retrospective analysis. *Journal of Chromatography A, 1389*, 133–138. https://doi.org/10.1016/j.chroma.2015.02.052

Hansson, A., Knych, H., Stanley, S., Thevis, M., Bondesson, U., & Hedeland, M. (2015). Characterization of equine urinary metabolites of selective androgen receptor modulators (SARMs) S1, S4 and S22 for doping control purposes. *Drug Testing and Analysis, 7*(8), 673–683. https://doi.org/10.1002/dta.1768

Jongedijk, E., Fifeik, M., Arrizabalaga-Larrañaga, A., Polzer, J., Blokland, M., & Sterk, S. (2023). Use of high-resolution mass spectrometry for veterinary drug multi-residue analysis. *Food Control, 145*, Article 109488. https://doi.org/10.1016/j.foodcont.2022.109488

Kaufmann, A. (2020). High-resolution mass spectrometry for bioanalytical applications: Is this the new gold standard? *Journal of Mass Spectrometry, 55*(9), Article e4533. https://doi.org/10.1002/jms.4533

Kaufmann, A., Butcher, P., Maden, K., Walker, S., & Widmer, M. (2017). Using in silico fragmentation to improve routine residue screening in complex matrices. *Journal of the American Society for Mass Spectrometry, 28*(12), 2705–2715. https://doi.org/10.1007/s13361-017-1800-2

Kaufmann, A., Butcher, P., Maden, K., Walker, S., & Widmer, M. (2022). Simplifying nontargeted analysis of PFAS in complex food matrixes. *Journal of AOAC International, 105*(5), 1280–1287. https://doi.org/10.1093/jaoacint/qsac071

Kohler, M., Thomas, A., Geyer, H., Petrou, M., Schänzer, W., & Thevis, M. (2010). Confiscated black market products and nutritional supplements with non-approved ingredients analyzed in the cologne doping control laboratory 2009. *Drug Testing and Analysis, 2*(11–12), 533–537. https://doi.org/10.1002/dta.186

Léon, A., Cariou, R., Hutinet, S., Hurel, J., Guitton, Y., Tixier, C., et al. (2019). HaloSeeker 1.0: A user-friendly software to highlight halogenated chemicals in nontargeted high-resolution mass spectrometry data sets. *Analytical Chemistry, 91*(5), 3500–3507. https://doi.org/10.1021/acs.analchem.8b05103

Polgár, L., García-Reyes, J. F., Fodor, P., Gyepes, A., Dernovics, M., Abrankó, L., … Molina-Díaz, A. (2012). Retrospective screening of relevant pesticide metabolites in food using liquid chromatography high resolution mass spectrometry and accurate-mass databases of parent molecules and diagnostic fragment ions. *Journal of Chromatography A, 1249*, 83–91. https://doi.org/10.1016/j.chroma.2012.05.097

de Rijke, E., Essers, M. L., Rijk, J. C. W., Thevis, M., Bovee, T. F. H., van Ginkel, L. A., et al. (2013). Selective androgen receptor modulators: In vitro and in vivo metabolism and analysis. *Food Additives & Contaminants: Part A, 30*(9), 1517–1526. https://doi.org/10.1080/19440049.2013.810346

Rojas, D., Dervilly-Pinel, G., Cesbron, N., Penot, M., Sydor, A., Prévost, S., et al. (2017). Selective androgen receptor modulators: Comparative excretion study of bicalutamide in bovine urine and faeces. *Drug Testing and Analysis, 9*(7), 1017–1025. https://doi.org/10.1002/dta.2113

Schmidt, K. S., & Mankertz, J. (2018). In-house validation of a liquid chromatography–tandem mass spectrometry method for the determination of selective androgen receptor modulators (SARMS) in bovine urine. *Food Additives & Contaminants: Part A, 35*(7), 1292–1304. https://doi.org/10.1080/19440049.2018.1471222

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., et al. (2014). Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environmental Science & Technology, 48*(4), 2097–2098. https://doi.org/10.1021/es5002105

Segal, S., Narayanan, R., & Dalton, J. T. (2006). Therapeutic potential of the SARMs: Revisiting the androgen receptor for drug discovery. *Expert Opinion on Investigational Drugs, 15*(4), 377–387. https://doi.org/10.1517/13543784.15.4.377

Steiner, D., Malachová, A., Sulyok, M., & Krska, R. (2021). Challenges and future directions in LC-MS-based multiclass method development for the quantification of food contaminants. *Analytical and Bioanalytical Chemistry, 413*(1), 25–34. https://doi.org/10.1007/s00216-020-03015-7

Temerdashev, A. Z., Azaryan, A. A., Labutin, A. V., Dikunets, M. A., Zvereva, I. O., Podol'skii, I. I., … Balabaev, I. A. (2017). Application of chromatography–mass spectrometry methods to the control of sport nutrition and medicines marketed via internet. *Journal of Analytical Chemistry, 72*(11), 1184–1192. https://doi.org/10.1134/S1061934817110090

Thevis, M., Gerace, E., Thomas, A., Beuck, S., Geyer, H., Schlörer, N., et al. (2010). Characterization of in vitro generated metabolites of the selective androgen receptor modulators S-22 and S-23 and in vivo comparison to post-administration canine urine specimens. *Drug Testing and Analysis, 2*(11–12), 589–598. https://doi.org/10.1002/dta.211

Van Wagoner, R. M., Eichner, A., Bhasin, S., Deuster, P. A., & Eichner, D. (2017). Chemical composition and labeling of substances marketed as selective androgen receptor modulators and sold via the internet. *JAMA, 318*(20), 2004–2010. https://doi.org/10.1001/jama.2017.17069

Ventura, E., Gadaj, A., Buckley, T., & Mooney, M. H. (2020). Development of a multi-residue high-throughput UHPLC–MS/MS method for routine monitoring of SARM compounds in equine and bovine blood. *Drug Testing and Analysis, 12*(9), 1373–1379. https://doi.org/10.1002/dta.2875

Ventura, E., Gadaj, A., Monteith, G., Ripoche, A., Healy, J., Botrè, F., … Mooney, M. H. (2019). Development and validation of a semi-quantitative ultra-high performance liquid chromatography-tandem mass spectrometry method for screening of selective androgen receptor modulators in urine. *Journal of Chromatography A, 1600*, 183–196. https://doi.org/10.1016/j.chroma.2019.04.050

Xu, Z., Li, J., Chen, A., Ma, X., & Yang, S. (2018). A new retrospective, multi-evidence veterinary drug screening method using drift tube ion mobility mass spectrometry. *Rapid Communications in Mass Spectrometry, 32*(14), 1141–1148. https://doi.org/10.1002/rcm.8154

Yan, X.-t., Zhang, Y., Zhou, Y., Li, G.-h., & Feng, X.-s. (2022). Technical overview of Orbitrap high resolution mass spectrometry and its application to the detection of small molecules in food (update since 2012). *Critical Reviews in Analytical Chemistry, 52*(3), 593–626. https://doi.org/10.1080/10408347.2020.1815168

Zomer, P., Mol, H., & Ur, R. W. (2020). Full-scan fragmentation options for the detection of food contaminants by an Affordable LC-Q-Orbitrap MS. *Brazilian Journal of Analytical Chemistry, 9*(34), 87–89.