SPECIAL GUEST EDITOR SECTION

European Analytical Criteria: Past, Present, and Future

Lynn Vanhaecke

Ghent University, Faculty of Veterinary Medicine, Department of Veterinary Public Health and Food Safety, Laboratory of Chemical Analysis, Salisburylaan 133, B-9820 Merelbeke, Belgium **PETRA GOWIK** Federal Office of Consumer Protection and Food Safety (BVL), EC Community Reference Laboratory for Residues, PO

Box 110260, D-10832 Berlin, Germany

BRUNO LE BIZEC

LABoratoire d'Etudes des Résidus et Contaminants dans les Aliments (LABERCA), ONIRIS, Atlanpôle de Chantrerie, Route de Gachet, BP 50707, 44307 Nantes Cedex 03, France

LEEN VAN GINKEL

RIKILT-Institute of Food Safety, European Community Reference Laboratory (CRL), PO Box 230, 6700 AE Wageningen, The Netherlands

EMMANUELLE BICHON

LABERCA, ONIRIS, Atlanpôle de Chantrerie, Route de Gachet, BP 50707, 44307 Nantes Cedex 03, France MARCO BLOKLAND

RIKILT-Institute of Food Safety, CRL, PO Box 230, 6700 AE Wageningen, The Netherlands

HUBERT F. DE BRABANDER

Ghent University, Faculty of Veterinary Medicine, Department of Veterinary Public Health and Food Safety, Laboratory of Chemical Analysis, Salisburylaan 133, B-9820 Merelbeke, Belgium

In this paper, the past, present, and (possible) future of the European analytical criteria for residues are described. The elaboration of the revision of Commission Decision 93/256/EC was a long process starting in 1996 and ending with the formation of a European Commission (EC) working group in 1998. This working group took account of developments in scientific and technical knowledge at that time and produced a draft version of the revision within 6 months. The revision, finally published in 2002 (2002/657/EC), includes new ideas on the identification of analytes and the criteria for performance assessment as well as validation procedures. Currently (2009), the evolution in analytical equipment and progress in scientific research, accompanied by recent European regulatory changes, demands an update or revision of the 2002/657/EC.

In the European Union (EU) the inspection of live animals and products of animal origin for the presence of residues of veterinary drugs and specific contaminants is regulated by the European Commission (EC; 1–3). According

to Council Directive 96/23/EC (2), two main groups of substances must be monitored to guarantee a high level of protection of human health in relation to its food from animal origin: group A and B substances. Group A comprises the prohibited growth-promoting agents and the prohibited substances for which no maximum residue limits (MRLs) could be established. Group B encompasses all registered veterinary drugs having an MRL and other contaminants, as summarized in Table 1. Since 1987, an extensive network of analytical residue laboratories has gradually been created in the EU for the purpose of veterinary inspections. This network consists of a hierarchical system of so-called Field Laboratories, National Reference Laboratories, and Community Reference Laboratories (CRLs). The quality criteria for the analysis of the above-mentioned residues and contaminants by these laboratories are described in a series of Commission Decisions (4-6). These decisions need to be revised on a regular basis to take into account the current scientific knowledge and the latest technical improvements.

In 1996, the Commission initiated a complete legal and technical revision of the two criteria decisions originating from 1993 (4, 5). Due to the complex nature of this revision process, the Commission in May 1998 designated a working group to draft new or revised criteria. This working group was chaired by François André (LABERCA, Nantes, France) and took account of developments in scientific and technical knowledge at that time. This revision was finally published in 2002 (6), and formed a regulatory basis for the performance of analytical methods, including guidelines on the identification of analytes, the interpretation of criteria for performance assessment, and

Guest edited as a special report on "Analytical Methods for the Determination of Veterinary Drug Residues in Food Products of Animal Origin" by Badar Shaikh.

Corresponding author's e-mail: Lynn.Vanhaecke@ugent.be

Table 1. Overview of Group A and Group B substances

Group A: Substances with anabolic effects and unauthorized substances

- · Stilbenes, stilbene derivatives, and their salts and esters
- · Antithyroid agents
- Steroids
- · Resorcylic acid lactones, including zeranol
- -Agonists
- Compounds included in Annex IV of Council Regulation (EEC) 2377/90 including *Aristocholia* spp. and preparations thereof, chloramphenicol, chloroform, chlorpromazine, colchicine, dapsone, dimetridazole, metronidazole, nitrofurans (including furazolidone), and ronidazole

Group B: Veterinary drugs and contaminants

- Antibacterial substances, including sulfonamides and quinolones
- Other veterinary drugs
- Antihelminthics

Anticoccidiostats, including nitroimidazoles

Carbamates and pyrethroids

Carbadox and olaquindox

Sedatives

Nonsteroidal anti-inflammatory drugs (NSAIDs) Other pharmacologically active substances

· Other substances and environmental contaminants

- Organochlorine compounds including PCBs
- Organophosphorus compounds
- Chemical elements
- Mycotoxins
- Dyes
- Others

the set-up of validation procedures. This final publication was preceded by publications in scientific journals (7) and in proceedings of several scientific conferences.

The goal of the 2002/657/EC (6), as well as of its predecessors, was the establishment of criteria for method assessment in the field of residue analysis, which was a modern approach in contrast to the use of standardized methods. Criteria Decision 93/256/EC (4) still differentiated between routine and reference methods, which was overcome with the 2002/657/EC (6). As a result, the criteria approach allowed the continuous application of new methods, analytical techniques, and analytical equipment, thereby allowing tracking of scientific developments more easily. Laboratories were able to apply their own methods and introduce new techniques. This also led to the introduction of in-house validation, i.e., the verification of the fitness for purpose of individual methods, which in turn requires the establishment of criteria for performance assessment. This

approach was first realized with Commission Decision 87/410/EC (8) and then continuously updated.

The elaboration of Decision 2002/657/EC (6) also introduced several new aspects distinct from its predecessors, with particular regard to the identification of residues or contaminants of interest and a complete new approach to validation. In the old legislation, the criteria, which were assessed for identification purposes, were the retention time, relative retention time, and spectra resulting from the use of different analytical techniques (UV and MS). When selected ion monitoring (low resolution) was applied, at least four ions with a particular ratio had to be present to claim a result as "positive." In reaction to this superseded legislation, the 2002/657/EC (6) introduced a system of identification points and laid down permitted tolerances for the relative intensities of the detected ions. Moreover, the result expressions "positive" and "negative" were replaced by "compliant" and "noncompliant" which fit better with the different substance groups under consideration in Council Directive 96/23/EC (2), Annex I. Indeed, MRL substances may be present (positive), but at concentrations below their respective MRL values, and thus "compliant."

Nevertheless, deeper considerations concerning the application of the criteria approach led to the problems of, first, how to give further advice with regard to positive declarations (when is a result "noncompliant"?) and, second, how to give reasonable advice on how to validate a method, since validation procedures considerably influence the quality of the results of the validation parameters. All members of the EC working group agreed that an exact recipe should be provided to reach a certain harmonization among the laboratories and a comparability of the individual methods performances. Therefore, some additional guidelines were published, e.g., Directorate General for Health and Consumer Affairs (SANCO)/2004/2726 (topics: recovery correction, use of QC samples, clarifications concerning validation, and parallel extrapolation); SANCO/2004/2726 rev. 1 (topics: measurement uncertainty and validation of methods for substances with sum-MRL); and SANCO/2004/2726 rev. 2 [topic: guidelines for the implementation of Commission Decision (CD) 2002/657/EC regarding some contaminants; 9].

The present Decision 2002/657/EC (6), however, has not been revised since its publication, despite the evolution in analytical instrumentation that has occurred with an unforeseen speed the last decade. In particular, the introduction of ultra-HPLC (U-HPLC) and high-resolution, accurate mass, full scan MS [e.g., time-of-flight (TOF) and Fourier transform (FT) MS] and the expanding use of GC coupled to combustion isotope ratio mass spectrometry (GC/C-IR-MS) highlight the need for new criteria to be developed and implemented. In addition, the Import Decision 2005/34/EC (10) and the new Regulation of the European Parliament (CD 470/2009; 11), respectively, introducing the concept of Reference Point for Action (RPA) and further formalizing this concept for all compounds for which administration is prohibited, stress the urgent need for a

Measuring technique	Substances Annex 96/23/EC	Limitations
HPLC or GC with MS detection	Groups A and B	Only if following either an on-line or an off-line chromatographic separation
		Only if full-scan techniques are used or using at least three (group B) or four (group A) identification points for techniques that do not record the full mass spectra
HPLC or GC with IR spectrometric detection	Groups A and B	Specific requirements for absorption in IR spectrometry have to be met
HPLC-full scan diode array detector (DAD)	Group B	Specific requirements for absorption in UV spectrometry have to be met
HPLC-fluorescence	Group B	Only for molecules that exhibit native fluorescence and molecules that exhibit fluorescence after either transformation or derivatization
Two dimensional (2-D) TLC-full scan UV-Vis	Group B	2-D HPTLC and co-chromatography are mandatory
GC-electron capture detection	Group B	Only if two columns of different polarity are used
HPLC-immunogram	Group B	Only if at least two different chromatographic systems or a second, independent detection method are used

Table 2. Suitable confirmatory methods for organic residues or contaminants

revision of the current criteria for performance assessment, identification, and validation in residue analysis. Therefore, in this review, some first suggestions are given towards the revision of 2002/657/EC (6) by proposing the potential alterations and additions that are requisite to include the aforementioned scientific, technical, and regulatory progresses, as well as the need for harmonization among laboratories of current analytical criteria and validation procedures for residue analysis.

Performance Criteria in Residue Analysis: U-HPLC/HR-MS

The current applicable performance criteria for residue analysis are described in Commission Decision 2002/657/EC (6). For banned (A) substances, the emphasis is on the identification of the substances in a large number of matrixes in a concentration as low as possible (zero tolerance principle). In this case, at first, screening methods targeting multiple residues need to be developed and, second, confirmatory methods. For B substances, methods for the quantitative screening and/or confirmation of substances only in edible matrixes need to be developed.

For screening purposes, 2002/657/EC (6) specifies that only those analytical techniques for which validity can be demonstrated and have a false compliant rate of <5% at the level of interest shall be used. High throughput and a low percentage of false compliant (false negative) results are the main criteria for screening methods, requiring a value for CC (detection capability) below the minimum required performance limit (MRPL) or MRL. The ideal screening method, therefore, combines these criteria with a low percentage of false noncompliant (false positive) results, but this reflects more the economics of laboratory testing and not the principle of screening, the only purpose being to sift out the "noncompliant" samples from a large population.

In the case of a suspected noncompliant result, this result shall be confirmed by a confirmatory method. In Table 2, the different methods or method combinations that, according to Commission Decision 2002/657/EC (6), are considered suitable as confirmatory for the identification of organic residues or contaminants, are presented. Confirmatory methods have to fulfill many requirements, including the typical parameters like trueness/recovery, precision, specificity, and application; the criteria for identification; and the newly introduced parameters CC (decision limit) and CC .

More specific requirements have been indeed included in Commission Decision 2002/657/EC (6) with regard to the chromatographic separation and MS detection, and a system of identification points was introduced to interpret the data for confirmation of group A and B substances. However, both the chromatography requirements as the criteria for analyte identification are absent or unsatisfactory for the latest analytical techniques and do not include the newest scientific findings. In addition, direct implications of the modern analytical techniques and regulatory actions concerning the RPA concept are not included in current performance criteria determinations, in particular with regard to the principle of the harmonized limit for residue control, i.e., CC

Identification Points

From recent review articles (12–15), it may be concluded that HPLC/triple quadrupole tandem MS (HPLC/QqQ-MS/MS) is currently the preferred method for residue analysis. The majority of current residue and contaminant analyses rely on the high sensitivity and

Function	Mass resolution ^a	Mass accuracy (mDa)	Remarks
Screening	10000	50	Relative retention time 2.5%
Confirmation	10000	5	1.5 IPs/ion or product-ion, minimum 1 ion ratio, relative retention time 2.5%
HR confirmation	20000	5	2 IPs/ion or product-ion, minimum 1 ion ratio, relative retention time 2.5%
MS/MS identification of unknowns	10000	5	Confirm postulated structure by NMR spectroscopy and/or confirm accurate masses at mass resolution 70 000 (FWHM)

Table 3. P	roposal for add	litional HPLC/MS	identification criter	ia to be suppleme	ented to the 2002	/657/EC (re	f. 6);
adapted fror	m Nielen et al. (ref. 20)					

^a Full width at half maximum (FWHM) at *m*/*z* 400.

selectivity of the selected reaction-monitoring (SRM) mode of QqQ-MS/MS. The two-stage mass selection enables the detection, identification, and quantification of preselected targets at low g/kg levels in complex biological matrixes, such as urine, feces, tissue, feed, and hair (12, 14).

As mentioned before, the Decision laid down permitted tolerances for the relative intensities of the detected ions, clearly specifying the maximum deviation between the observed and the expected ion ratios. This use of relative ion abundance tolerance windows for ion abundance ratios in the mass spectra of particular chemicals makes good sense in theory (e.g., 30% RSD for an average relative ion abundance of 20% yields an ion tolerance window of 14-26% for achievement of an acceptable identification result). In practice, this ever tighter ion ratio for least intense ions might make "acceptable" identification more difficult in complex matrixes, especially at lower concentrations. Another practical issue is the fact that different ways to measure ion ratios might give somewhat different results. A most appropriate way to determine ion ratios in SRM should, therefore, be devised, i.e., the integration of analyte peaks from selected ion chromatograms using different ions. Dividing the integrated peak areas of each ion by the integrated base peak corresponds to the average relative ion abundance.

In addition to the relative ion abundance window, the 2002/657/EC (6) also introduced a system of identification points (IPs) for MS detection. For the confirmation of Group A and B substances, a minimum of four, respectively three IPs, are required. When using the SRM approach, this implies four identification points can be collected by obtaining two transitions with 1.5 identification points each and one point for the precursor. In some cases, however, false noncompliant results may be obtained by this approach (16). Indeed, the Commission Decision does not discuss the relative importance of certain transitions. It is widely accepted that losses of water and carbon dioxide do not provide the same selectivity as some other more unique neutral losses. It, however, does discuss the origin of the selected diagnostic ions, i.e., originating not from the same part of the molecule (6), a criterion that, in practice, is very difficult to monitor. Moreover, it must be taken into account that the quality of a confirmation based on transitions produced by low-mass precursor ions is, in general, less than that based on high-mass precursor ions. In particular, the presence of coeluting endogenous isobaric compounds at relatively high concentrations might pose a serious risk in this situation. Although there is a small likelihood that this will happen with the more commonly analyzed matrixes, this risk increases with matrixes for which the analyte has not been explicitly validated. Under these circumstances, findings should be confirmed by a third transition (if possible) by co-chromatography and by analyzing additional blank samples of the same or a similar matrix whenever possible as described in Chapter 2.3.1 (6), or even more advantageously by an orthogonal criterion like full-scan accurate mass-based techniques.

The use of full-scan MS approaches (e.g., TOF, quadrupole TOF, FT ion cyclotron resonance, or FT Orbitrap) has increased tremendously in recent years. Full-scan MS approaches offer the possibility to simultaneously analyze a virtually unlimited number of compounds. Furthermore, the retrospective "post-targeted" evaluation of old data offers the possibility to detect non-"a priori" selected analytes. Moreover, their accurate-mass capabilities support the reconstruction of highly selective, accurate-mass chromatograms of target residues in complex matrixes. To allow the detection of residues in the low ppb (g/kg) or ppt (ng/kg) concentration ranges, which is required by legislation, very sensitive full-scan analyzers are required. The medium resolution of TOF systems significantly affects the selectivity, and therefore, the sensitivity gain, compared to unit-resolution scanning MS. The utility of high-performance TOF-MS application has been demonstrated for the multicompound screening of veterinary drugs in different matrixes of animal origin (17, 18), and for doping agents in human urine (19) and pesticide residues in crops (20). It should, however, be noted that accurate-mass determination without proper

mass-resolution criteria might lead to false compliant results, both in MS screening and MS/MS confirmation. In this context, the high resolving power (up to 100 000) of FT MS technologies provides mass accuracy (below 2 ppm) resulting in both high selectivity and sensitivity for complex sample residue analysis.

The 2002/657/EC (6) defines high-resolution MS (HRMS) as MS at a mass resolution of 10 000 according to the 10% valley definition. This resolution of 10 000 according to the 10% valley rule corresponds to a resolution of 20 000 FWHM (full-width at half maximum) for modern instruments (e.g., (Q)TOF-MS, FT ion cyclotron resonance-MS, and FT Orbitrap-MS). The criteria for HRMS are as follows: 2.0 IPs are earned for each precursor ion, and 2.5 IPs for each product ion. No criteria for mass accuracy are described. Several studies have, however, demonstrated that false compliant (false negative) results can be obtained when the mass resolving power of the MS is insufficient to separate analyte ions from isobaric coeluting sample matrix ions (15, 21). Therefore, a proposal of additional HPLC/MS criteria, for which further discussions with regard to their inclusion in 2002/657/EC (6) are required, is presented in Table 3.

This proposal acknowledges the performance of state-of-the-art HPLC/(Q)TOF-MS instruments by ascribing 1.5 IPs/ion versus 1.0 in conventional low-resolution HPLC/MS, but it does not allow the use of only one parent and product ion, which should remain the exclusive domain of HRMS. It should, however, be noted that the original purpose of HRMS was to detect the exact mass of the parent ion and not to produce fragments. Modern Q-TOF and standalone Orbitrap or LTQ-Orbitrap instruments do, however, allow the acquisition of, respectively, tandem or multiple MS mass spectra, enabling the use of product ions in medium or high resolution. For the identification of metabolites or unknown compounds, a higher resolution is, however, warranted when NMR spectra cannot be obtained. According to Nielen et al. (21), the proposed resolution of 70 000 (FWHM) will ensure that reliable elemental compositions of product ions differing in one CO, C₂H₄, or N₂ substructure can be obtained up to m/z 400. In some cases, a resolution of 100 000 may even be necessary to achieve discrimination of product ions from matrix interferences, in particular when utilizing more generic sample preparation procedures and multiresidue procedures.

According to the 2002/657/EC (6), MS methods may be used only as confirmatory methods after chromatographic separation (offline or online). For HPLC/MS, the 2002/657/EC (6) decision states that suitable HPLC columns should be used. However, there is no appropriate definition of what "suitable" is. The user currently earns the same number of identification points regardless of whether an analyte leaves a low-resolving column at a relatively low k (capacity factor) value or if it is well resolved by a high-resolving U-HPLC column. Indeed, the introduction of pressure-stable 1.7 m particulate packing materials and novel low-dead-volume, high-pressure (to 10^5 kPa) HPLC equipment (U-HPLC) provides strategies to improve resolution while maintaining or even shortening run times. Therefore, the time has come to consider the use of 1.7 m, or at least 3 m, packed columns whenever possible for confirmation or provide an additional confirmatory power for higher k values by means of granting more identification points. Also, the tolerance of the correspondence of the relative retention time of the analyte to that of the calibration solution (now 2.5% for classical HPLC) should be revised based on existing knowledge of the reproducibility of the retention times in U-HPLC.

MRL and MRPL/RPA versus CC and CC

When can a "positive" result of an investigated sample be appointed as "noncompliant"? To answer this question at first the legal framework under which the test is performed, e.g., with respect to the classification "Group A" substance (banned substance) or "Group B" substance (substance for which a legal MRL has been established) needs to be addressed. Secondly, information regarding the precision of the measurement needs to be available. Furthermore, in special cases, knowledge of a possible background (natural or endogenous) level is necessary.

The knowledge of the precision at the "level of interest" forms the basis for further considerations on how to assess a noncompliant result. Therefore, CD 2002/657/EC requires validation at this level (6). For B substances, this level of interest clearly is the MRL. However, for methods to control banned substances (Group A) for which no safe level has been established and, subsequently, usually with zero-tolerance, the term MRPL was established. It was intended as a level at which every official laboratory in residue control had to be able to validate, i.e., a safe level, but a tool for harmonization between laboratories. MRPLs are formally established for four substances (substance groups), namely, chloramphenicol, medroxyprogesterone acetate, and nitrofuran metabolites (22), as well as malachite green and leuko malachite green (23). These MRPLs are defined for particular analyte/matrix combinations. For other A substances and for nonauthorized Group B substances or for substances for which no MRL was established for particular matrixes, MRPLs have been proposed by the CRLs and, in the meantime, called recommended concentrations. These combined proposals have been published as a CRL Guidance Paper in 2007 and are available through the various CRL residue websites (24). Since the first definition of the MRPL in 2002 and the subsequent setting of various MRPL values, events took a different turn when Import Decision 2005/34/EC, Art. 2 (10) was published in 2005. This Decision introduced the concept of using MRPL values as RPAs for imported goods. The RPA is defined as the concentration above which further action is required, i.e., destruction or redispatching. However, if the analyte of interest is detected below the RPA but above the CC, the duty to report the result is still valid, and in the case of a recurrent pattern, the Commission may take other actions. Recently, the new Regulation of the European Parliament and of the Council (470/2009) further formalized this concept for all compounds for which administration is prohibited (11) and enlarged the approach to the European single market. There are at least mixed feelings on this development among the scientific community inside and outside the EU. The main problem is the fact that when banned substances are detected and their identity fully confirmed, but the determination of the mass concentration results in a value between the CC and RPA, the assignment (or particular animal) is considered "safe" for consumers, and the implications for the responsible party are limited. To circumvent this problem to a large degree, the RPA should be set on the basis of the lowest residue concentration. Therefore, this lowest residue concentration should be quantified with an analytical method validated in accordance with Commission requirements. Validation in accordance with CD 2002/657/EC requires a vast amount of work. As a result, the minimization of the concentration range between CC and the RPA, i.e., the difference between LOD (3 S/N) and LOQ (6 or 10 S/N), should be achieved. Nevertheless, this implies that the CC is still the limit at or above which a result is analytically noncompliant, provided the identification criteria are fulfilled in confirmatory analysis (6). The term CC denotes a critical concentration concerning the or producer error. This CC should be lower than the MRPL (and in the future the RPA). The CC concerns the or consumer error and should be lower than or equal to the MRPL (9) and, consequently, by definition lower than or equal to the RPA.

As mentioned before, prior to confirmatory analysis, screening methods need to be developed as a first step to analyze large numbers of samples and to exclude all compliant samples from further, more advanced testing. These methods are characterized mainly by CC (also known as the consumer's error or the possibility of getting a "false negative"result). Therefore, CD 2002/657 (6) defines CC for substances for which no permitted limit has been established as the lowest concentration at which a method is able to detect truly contaminated samples with a statistical certainty of (1-). The practical determination of this value depends on the type of method used, i.e., quantitative or qualitative. Quantitative methods provide a measurable signal relating the intensity directly to the concentration. CD 2002/657 (6) gives two alternative approaches based on either the calibration curve or on the analyses of a set of 20 blank materials fortified at the level of the decision limit (to be determined first). This approach is applicable for both A and B substances. A special case is the use of qualitative methods providing a yes/no answer without a clearly quantifiable response. Here the process, in principle, is iterative. A set of 20 blank samples is fortified at levels just above the decision limit. The concentration at which 19 out of 20 samples are identified correctly as "noncompliant" is considered to be the best estimate of CC . This approach is nowadays also used in HRMS when no quantifiable blank response is present. For screening methods, CC is always higher than the CC . A special case not mentioned in the EU legislation, but frequently observed in residue laboratories, is that of semiquantitative methods. Such methods are in principle quantitative methods, e.g., based on chromatography combined with MS in the presence of an isotope-enriched

internal standard among many others. Moreover, in multiresidue methods, it is not uncommon to have a procedure that is quantitative for one analyte and semiquantitative for a second analyte. In some fields, such methods are used as "definitive methods" for metrological purposes, providing the best estimate of the true concentration. However, in residue analyses the situation is different. The combination of very low concentrations with complex biological matrixes with a variety of stability and isolation problems makes it very difficult to provide data with good precision. Many authors, therefore, only claim their method to be semiquantitative. For screening methods in practice this usually does not impose any problems.

For confirmatory methods the situation is more complicated. Compounds with a permitted level need the following statement to be confirmed: Does the concentration of the substance with an identity identical to that of the banned substance tested for truly exceed the tolerance level (MRL)? For compounds with no permitted level the only statement to be confirmed is: Is the identity of the substance detected truly identical to the identity of the banned substance tested for? Principally there is no quantitative aspect in this other than that the concentration should exceed CC . The confirmatory method, therefore, will focus on collecting the necessary (four) identification points and puts less emphasis on the determination of the concentration. For quantitative methods it usually will be possible to provide sufficient insight into the actual concentration to state that it exceeds CC . This is not self-evident, since if the true concentration equals CC, this will result in an analytical result higher than the CC in only 50% (Gaussian distribution) of all cases. In such cases it becomes a real option that the result of the confirmation will be that there is insufficient information to conclude that the concentration exceeds CC , which may lead to a complicated situation in court. However, for qualitative methods it is very difficult to determine CC . The only feasible option is to determine CC by the iterative procedure described above for the weakest signal to be monitored. Subsequently, it can be verified whether or not all the other (stronger) signals are indeed present and provide the four identification points (correct ion ratios).

Performance Criteria for GC/C-IR-MS in Steroid Analysis

Since the introduction of the European hormone ban in 1986, its implementation has proved to be an enormous challenge to regulatory authorities. In an effort to challenge the black market trade in hormones, there have been many analytical advances over the past few decades (25). In recent times, improved MS-based confirmatory analysis has greatly increased the likelihood of detecting xenosteroid abuse. There are many anabolic steroids produced by the animal body that enhance growth performance. Beyond designer steroids, doping with steroids naturally produced by the animal body is a challenge to detection. Indeed, urinary testosterone or estradiol concentrations are an unsuitable metric because of the large range of inter-individual urinary steroid excretion, eliminating the possibility that natural steroid hormone administration could be detected on the basis of concentration (26). Successful analytical strategies have been already reported for hair samples focusing on natural steroid esters residues (27-29), or specific conjugates in different organs (30, 31) or fluids (32, 33). The origin of natural steroids, endogenous versus exogenous, may be also determined in cattle urine on the basis of their isotopic composition (34, 35). The current confirmatory approaches are mainly based on the ${}^{13}C/{}^{12}C$ ratio determination by GC/C-IR-MS. Indeed, the administration of a synthetic steroid to an animal leads to the alteration of the ${}^{13}C/{}^{12}C$ ratio of excreted steroid metabolites, whereas their endogenous precursors remain unchanged (known as endogenous reference compounds, or ERC). Different analytical methods have been developed in Europe but rarely used for official control of testosterone (36-38), estradiol (39), and cortisol (40). GC/C-IR-MS is not discussed in the 2002/657/EC Decision (6) due to its relatively recent introduction in the field, and to the very limited number of laboratories competent regarding the technology and its application for trace characterization of steroids. Robust measurement and nonambiguous data must be produced; for that reason, several analytical criteria, GC/C-IR-MS-dependent and specific, have to be fulfilled, and some precautions have to be taken before any conclusion regarding the compliance/noncompliance of the sample. In these conditions, the GC/C-IR-MS technique and the use of

¹³C (carbon isotope ratio) values will be recognized as the gold standard in differentiating exogenous from endogenous steroids in urine in breeding animals, as it is already in the antidoping field (41, 42).

Nonetheless, the high precision of these ¹³C/¹²C measurements does not always lead to highly accurate results (43). The problem depends upon several factors: steroid extraction and purification (quality and efficiency); steroid derivatization (free versus derivatized); and instrumentation (chromatographic separation, sufficient mass spectrometric sensitivity to detect the changes in the ¹³C values and the low concentrations in some biological samples, and kinetic and/or enzymatic fractionation). Furthermore, ambiguities remain regarding the baseline ¹³C values used in establishing reference ranges. Explanations are linked to diet, metabolic fractionation, and physio-pathological variations among the tested animal population.

Accuracy of the GC/C-IR-MS Measurement

The accuracy is generally determined during the validation process and checked systematically during each run. Repeated reference CO_2 pulses may be flushed into the source. The ratio ${}^{13}C/{}^{12}C$ is measured for the reference peaks, and the SD is then determined to be within a laboratory limit. Because isotopic fractionation is a potential source of inaccuracy, the whole analytical process has to be checked. It is generally performed by adding at least one surrogate sample in each sequence, and by comparing the ${}^{13}C/{}^{12}C$ value of each added steroid with the

corresponding reference (already characterized by an independent combustion ¹³C analysis). Other quality controls should be analyzed with each sequence to demonstrate that the procedure is capable of identifying a positive urine sample that contains ¹³C-depleted steroid peaks, as well a negative sample that does not show a significant ¹³C/¹²C difference between metabolites and ERC. Stability and reproducibility of GC/C-IR-MS measurements can be monitored by determining the ¹³C value of an internal standard in each sample. This ¹³C value should lie within an estimated uncertainty from the ¹³C value measured by independent combustion ¹³C analysis. The ¹³C value calculated by the software is relative to the 13 C value of the CO₂ or steroid reference materials that have been determined by comparison to a certified standard such as the NBS-19 described by the International Atomic Energy Agency to represent Vienna Pee Dee Belemnite (VPDB; 43).

Specificity of the Final CO₂ Peak

The operator has to demonstrate that the integrated signal is exclusively coming from the target compound. In other words, it is necessary to verify that there are no coeluting interferences coming either from the biological matrix or from the analytical instrument, e.g., oil or stationary phase. The general strategy used to control the so-called "purity" relies on the comparison of mass spectra between the target analyte (in the sample) and the corresponding reference (either standard in the same sequence or certified library). It is performed with a second injection of the extract on a conventional GC/MS system or in the same acquisition when the instrument is equipped with a T-junction. A purity score not significantly different from 100% is expected.

Identification of the Target Compounds

Two main parameters are used for identification, i.e., retention time and ion ratio. The chromatographic criterion already defined in the 2002/657/EC decision must be directly applied (i.e., 0.5% for relative retention time in GC). This is the case as well for identification criteria; double injection in GC/MS is necessary for that purpose. A solvent blank that does not contain any steroid and/or unknown compounds should be the first injection of a sequence to verify that the system is void of contamination.

Noncompliance Criteria

The analytical and physiological (intraindividual and interindividual) variability should be taken into consideration to determine the noncompliance criteria. If the difference between the ${}^{13}C_{VPDB}$ values of an anabolic steroid or its metabolite and the ERC exceeds a given limit, this is considered evidence for the presence of exogenous steroids. In the antidoping field, an official threshold has been set up and published by the World Antidoping Agency (44) "...the results will be reported as consistent with the administration of a steroid when the ${}^{13}C/{}^{12}C$ value measured for the metabolite(s) differs significantly, i.e., by 3 delta units or more from that of the urinary reference steroid chosen... ." No



Figure 1. Linear (a) versus worst case (b) extrapolation.

official criterion regarding this difference is available in the food safety area. Norms and extremes of steroid isotope ratios have still to be validated on a wider population. These data will form a basis against which future natural hormone fraud testing can be compared to more accurately distinguish clean animals from those that were treated. Samples should be reported as consistent with the administration of testosterone/estradiol when the ¹³C/¹²C values measured for the metabolite(s) differ by three delta units or more from that of the reference steroid chosen.

Validation in Residue Analysis

Worldwide organizations, legislators, and scientists are working on principles and procedures for the validation of analytical methods (45-54). Method validation is very often based on ring trials (collaborative trials) for the determination of reproducibility, repeatability, and recovery. For quantitative methods this is mostly performed in a more or less harmonized way, based on the International Organization for Standardization (ISO) 5725 series, although even here more than one possibility exists on how to plan and conduct a collaborative trial (45). For qualitative methods, no such ISO Standard equivalent exists. Instead, organizations like AOAC INTERNATIONAL (46) and Nordic Committee on Food Analysis (NMKL; 52) have elaborated procedures and instructions on how to perform validations in this case, but they are restricted to microbiological methods. However, the principles were taken from ISO Standard 16140, which provides procedures for the validation of qualitative and quantitative microbiological methods but which, up to now, only refers to alternative, proprietary methods (8).

Meanwhile, more and more efforts have been made to lay down prescriptions for in-house validations that are applicable to certain types of methods. The European residue control of growth promoters and veterinary drugs in food-producing animals has played a leading role in this context. Already in the 1980s, criteria for the in-house validation of analytical methods were laid down legally (1). In the meantime, the Codex Alimentarius Committees for Veterinary Drugs in Food (CCRVDF; 48) and for Pesticide Residues (CCPR; 46), as well as standardization organizations like ISO, European Committee for Standardization (CEN), and AOAC (49-51) have made efforts to establish such prescriptions or have already done so. In-house validations are an important approach to validation insofar as, on the one hand, they lead to a verified method much more quickly, and on the other hand, they allow comprehensive checking of the proficiency of the method in the laboratory itself. This is often neglected after the transfer of a ring-trial-validated method or standard. So-called verification studies are often not performed invoking that, after all, a validated method has been used. This leads to the not to be underestimated danger of an unnoticed underperformance of individual laboratories, which is certainly not in the interest of consumer protection.

The stipulations described earlier concerning performance criteria of analytical methods together with the identification point system are based on the verifiability of quantitative and qualitative measurements or the degree of certainty of the identification of an analyte. These criteria and parameters are the ones that, in particular, need to be determined and checked (55) with the help of the detailed prescriptions for the validation of methods laid down in CD 2002/657/EC, Chapter 3 (6).

The harmonization of validation procedures was the principle aim of Chapter 3 of CD 2002/657/EC (6). Different validation procedures lead to different degrees of meaningfulness of the individual validation parameters, with the consequences that a comparability of laboratory and method performances cannot be achieved and a harmonized control is not possible. Therefore, there were considerations to solve the situation to a certain extent by providing very detailed prescriptions for in-house validations to be fulfilled by each official control laboratory, and to simultaneously include factors that have to be considered, like major and minor changes. Moreover, to ensure the real fitness for purpose of the applied methods, the concentration levels of interest were defined exactly, considering that the methods have to determine noncompliance. This means that the methods do not have to be validated necessarily at the LOD or LOQ, but at the level of interest. This level of interest might be the zero tolerance in the case of prohibited substances, but it is also very often a maximum legal limit like MRL or MRPL or an agreed-upon level like recommended concentrations. As a compromise among European scientists and legislators, two principally different validation procedures could be agreed upon: one according to a classical approach, and another one according to the, at the time, completely new alternative approach based on a statistical factorial experimental design (56, 57). Both approaches represent different paradigms and take into account different uncertainty components, the most significant of which is the uncertainty component resulting from the matrix, operator, environmental, and design parameters, among others, subsumed under the term of matrix mismatch component. Allowing these two approaches, it is inherently accepted that validation procedures deliver different results for the performance parameters and different certainties of the determined validation parameters (9). One example of different procedures within the classical validation is demonstrated in Figure 1. It can easily be derived that completely different results for CC are the case although the same set of data was used. Meanwhile, guidelines of SANCO 2726/2004 (9) have recommended the verification of CC by means of blank material spiked at the CC level calculated in that way in order to prove that the method is able to detect and to identify the analyte and to fulfill all criteria required by CD 2002/657/EC (6). Even more possibilities to determine are allowed by CD 2002/657/EC (6), e.g., the CC determination using the blank-noise procedure, which in the case of the most advanced triple-quadrupole or high-resolution MS techniques, leads to nearly no background noise and, therefore, too-low CC values.

In the meantime, the SANCO 2726/2004 (9) guideline has been published to support CD 2002/657/EC (6) with regard to some "questionable interpretations." In addition to the different possibilities offered by the Decision itself, the prescriptions are subject to numerous different interpretations by the analysts. For example, the validation is not always performed at the level of interest. Instead, analytes are validated at the zero level, although an MRL exists (unpublished proficiency test reports). Furthermore, the terms CC and CC are not always used in the proper sense of their meaning. Discussions have repeatedly shown that CC and CC are used in the same sense as LOD and LOQ, which is not correct. The CC is the critical concentration concerning the -error (56) and serves as a limit of decision at which a method is capable of determining the compliance with a limit (MRL, zero tolerance), taking into account the extended measurement uncertainty. The CC refers to the -error (false-negative error) and gives the limit at which a method is able to detect really contaminated samples with a certainty of,

e.g., 95%. Therefore, no correlation exists to the LOQ, which refers to the -error.

The term "major change," which is used in Chapter 3.1.2 of the Decision, also leaves a lot of room for interpretation. Its original meaning was that factors that become known in the course of the validation or method establishment, which may influence the measurement result or which are strongly reckoned to have a considerable influence on the measurement value, are to be validated separately in the framework of the classical validation. The reason for this is that the classical validation per se, i.e., due to the statistical concept it is based upon, is not able to separate influencing factors retrospectively. Therefore, separate validation studies are required. The definition of "major change" depends on the respective method and analyte and can, aside from suggestions or experiences from other laboratories, only be defined by the concerned analysts themselves. Frequently, for example, the matrix is a major change, but also different species as well as different operators, the time, and different equipment may have a considerable impact on the measurement values. This has to be assessed by the respective scientists and cannot be answered generally.

The measurement uncertainty and its determination have been discussed intensively for some years. The question is which measurement uncertainty components have to be included. This depends on the purpose of the analysis. CD 2002/657/EC (6) offers a very pragmatic solution for this problem, which has already been recognized by the European Analytical Chemistry network (EURACHEM; 58), by including the expanded measurement uncertainty in CC. CC is calculated using the within-laboratory reproducibility. The within-laboratory reproducibility, as defined in CD 2002/657/EC (6), already includes the matrix-mismatch uncertainty components, so that uncertainty components as they are inherent to every new sample through their specific nature and matrix as well as by way of time factors are also contained in CC . The measurement uncertainty determined on that basis offers a high safety factor regarding false positive decisions, but also delivers higher concentration levels for the limits of decision and the detection capabilities than would be the case when applying the conditions of the ISO standards (59). Being based on a statistical model that does not account for matrix-mismatch uncertainty components but for repeatability conditions, the application of ISO standards automatically leads to lower uncertainty ranges. This example shows again clearly that validation parameters determined by means of these different concepts are not comparable.

The explicit prescriptions on validation currently refer to confirmatory methods only. However, since the need for prescriptions for screening procedures was also expressed, a working group composed of representatives of CRLs is working on the establishment of a concept for the validation of screening methods. In this context, most problems can be seen in connection with unspecific procedures with high numbers of analytes. In these cases, the extent of the validation is extraordinarily large since each analyte has to be validated separately, which leads to a multiplication of the number of samples and, thus, of the required time for validation. Obviously, it is not possible for each routine laboratory to accomplish a complete validation. Concepts are to be developed allowing that the comprehensive, complete validation is carried out by a specialized laboratory once, and would then, provided that the same method is performed on the same equipment, only have to be verified for selected analytes and matrices in the other laboratories. As efficient as this procedure looks, its disadvantages become clear at first sight. It means a return, more or less, from the criteria approach, i.e., each laboratory can use and validate its own method, to the standardization principle, i.e., a specific method is to be used and no variations are allowed. Since both ways will remain open, it is up to every analyst to weigh the advantages and disadvantages of each approach.

The quintessential aim, namely, the harmonization of residue control, or of the assessment of positive results, has not been reached to the extent that seems desirable. Nevertheless, the control of veterinary drugs and growth promoting substances in food-producing animals is an area that has dealt with this question most exhaustively. The statistical backgrounds of criteria and validation parameters have been looked into intensively, and the required efforts connected with validation studies have been faced comprehensively. An important success of CD 2002/657/EC (6) and the discussions thereof can be seen in the high and still growing consciousness with regard to the difficulties with which residue control at trace levels is confronted. It would be gratifying if this consciousness was also extended to other (trace) analytical areas.

Criteria and Quality Control for Metabolomics

The general principle of metabolomics is to characterize biological samples through a large-scale generation of descriptors referring to chemical species present in these samples and accessible for analysis (60). MS today appears to be the technique of choice for such metabolic profiling studies (24, 61, 62), mainly because of its incomparable sensitivity (detection of minor but potentially informative metabolites) and specificity (structural elucidation purposes) as opposed to other techniques. The last generation of high resolution mass analyzers, including orbital trap (63-65) and TOF instruments (66, 67), especially appear well adapted for MS-based metabolomics by combining sensitivity in the full-scan mode, good dynamic range, high mass resolution, and high mass accuracy. Because of the relative youthfulness of these approaches, many efforts have been devoted to developments, testing, and comparisons in the last few years, usually leading to self- and tailor-made global strategies (68) from the sample preparation to the statistical analysis of the produced data, including extremely diverse analytical tools and bioinformatic software solutions (69-72). However, after this "first age" of metabolomics, one can observe today a further need for improved QC procedures and new analytical criteria in this field. And this requirement appears necessary not only to ensure the quality of the data/results produced, but also in the perspective of applying metabolomics for regulatory control purposes.

From a purely analytical point of view, MS-based metabolomics can be considered as a large multiendpoint measurement process and an extended multiresidue monitoring method. Consequently, some basic and classical QC procedures routinely applied for conventional multiparametric measurement methods can be directly transposed to metabolomics. The addition of several reference substances to each analyzed sample before any treatment may be, for instance, strongly suggested. These multiple internal standards permit an efficient monitoring, if not correction, of different analytical criteria such as retention time, signal abundance, or mass accuracy (73).

Chromatographic Separation

Because the global repeatability of the metabolomic process immediately appears as a key issue in this field, the temporal robustness of the chromatographic system is a first parameter to control either in HPLC or GC. Indeed, even if some very efficient software algorithms exist for chromatographic peak alignment (74-77), a very high degree of reproducibility in terms of retention time across various samples and injected sequences (i.e., intraday and interday) undoubtedly facilitates further data processing and the direct comparison between individual samples, for example in case of a posteriori verification of the presence or absence of a given signal. The use of multiple internal standards in each analyzed sample may serve as a first indicator of this chromatographic separation robustness, for example, by plotting the observed retention time of each reference compound in a control chart. Then a significant drift of this parameter (classically when out of the limit: mean 2 SD, where mean and its variability were determined during method development) may imply that a periodic or total cleanup of the chromatographic system is required, if not a complete replacement of the analytical column. But if this action usually appears as sufficient for conventional targeted monitoring, an additional careful examination of each entire total ion chromatogram appears justified in metabolomics. Fixing a defined criterion regarding this critical examination by the analyst remains quite difficult. Indeed, a progressive drift in terms of baseline, the presence of one or more atypical sample profiles within the analyzed batch, or the existence of chromatographic regions subject to ion suppression may directly impact the final results.

MS Fingerprinting

The repeatability of the MS metabolic fingerprinting is once again of prime importance. Systematic examination of the reference compounds added to each sample may also be of great value at this stage. In the same way as previously described for the retention time, monitoring the signal abundances observed for each internal standard across the analyzed samples through a control chart may reveal any drift of the system in terms of sensitivity. Because many applications of metabolomics are focusing on the identification of potential biomarkers, the concept of mass accuracy is of high importance, considering that all further structural elucidation hypotheses will be based on the assumption that the MW of the parent compound was determined sufficiently precisely. Therefore, TOF and orbital trap instruments used for metabolomics have to be very regularly calibrated with scrutiny in order to maintain a high degree of mass accuracy during all experiments, and most of all a good repeatability of this parameter (a systematic but constant mass error can be relatively easily corrected), once again to allow a posteriori verification of presence/absence of a given signal in previously analyzed samples. This parameter appears even more crucial in metabolomics than resolution, stressing again the need for introduction of new performance criteria (absolute target value and maximal authorized variability) and/or identification points (see earlier) regarding mass accuracy should be suggested. Last, but not least, a systematic decomposition of the observed total variability for the generated MS metabolic fingerprint may be proposed. Indeed, it could be of valuable interest to propose a standard operation procedure in terms of biological, analytical, and instrumental replicates. The inclusion of a defined number (n = 3 to 6) of replicates, either in terms of injections of each prepared sample extract or in terms of each sample to be analyzed, may permit a comprehensive evaluation of the main sources of signal variability, i.e., expected (due to the hypothesis of differences existing between subgroups of samples), biological (due to genetic polymorphism and many other environmental factors), analytical (due to the sample preparation procedure), and instrumental (due to the measurement system). Consequently, a multiple analysis of variance performed on all MS signals constituting the generated metabolic fingerprints may permit estimation for each component of this variability, with some maximal values to be designated as indicating acceptable analytical and instrumental quality. Because of the very high number of detected signals (typically between 1000 and more than 4000), this evaluation should be performed by considering not each signal individually but the general distribution frequency of all signals together, i.e., grouping the detected signals into several classes according to their observed variability. Then, the principle would be to determine the mean variability observed for the majority of the detected signals, and to compare this value to a proposed maximal limit (for example, 30-40%).

Post-Metabolomics

If the final objective of the developed metabolomic approach is to identify indirect biomarkers revealing the exposure to a chemical as a means for screening in the frame of regulatory analysis, these biomarkers will be more conventionally monitored in routine analysis upon their identification using the metabolomic approach. Indeed, from the moment these biomarkers of interest are known, it can be envisaged to measure them by simple or multistage MS with an appropriate quantification procedure. The question is, however, to fix some analytical criteria in this context, considering the concept of IPs introduced at the European level by the 2002/657/EC Decision (6) that is currently based on the unambiguous identification of specified compounds (parent drug or metabolite) explicitly included in official reference documents [for example, the 96/23/EC Directive for drug residues (2)]. Moreover, the monitoring of indirect biomarkers is currently not envisaged for screening. This means that some analytical criteria will have to be defined at this level, for example, considering one identification point/one unambiguously identified biomarker. Similar to the case for conventional methods, the concepts of CC and CC may also be transposed to each signal corresponding to these indirect biomarkers. However, besides these theoretical considerations, the road to reach this ultimate goal of metabolomics is certainly quite long, and in parallel to these analytical criteria issues, the question of how to validate properly such approaches will also be a major challenge for analysts and regulatory authorities.

Conclusions

In this paper, the history and present status of the performance and validation criteria for the analysis of residues in laboratories of the EU are reported. In addition, some first suggestions towards the revision of the current 2002/657/EC are proposed. Indeed, the progress of the last 10 years since the Decision has come into force makes a revision and enlargement of the criteria and prescriptions urgently necessary. The evolution in analytical instrumentation has taken place at an unforeseen speed in the last decade and, without any doubt, will continue to do so in the future. Moreover, the European legislation is an ongoing changing process. Therefore, the performance and validation criteria will have to be revised on a regular basis (e.g., each 5 years). Moreover, additional criteria for new techniques, such as GC/C-IR-MS, and emerging concepts such as metabolomics will need to be foreseen in the current revision as well.

Some elements of the 2002/657/EC, including the system of IPs and the in-house validation, are intended to be timeless but will require fine-tuning over the years. The prescriptions on validation of CD 2002/657/EC are indeed the most modern regulations existing in the framework of legal regulations or standards. Of course, critical readers and users will always find points that seem to need improvement, which the authors do not deny. Keeping analytical criteria up to date will be a never-ending story for the present and future generation of analysts.

Acknowledgments

We wish to thank all colleagues who contributed to the elaboration, development, interpretation, and discussions on CD 2002/657/EC. They are so numerous that they cannot all be named. The EC is also gratefully acknowledged for its support.

References

- Council Directive 86/469/EC of 16 September 1986 (1986)
 Off. J. Eur. Commun. 275, 36–45
- (2) Council Directive 96/23/EC of 29 April 1996 (1996) *Off. J. Eur. Commun.* **125**, 10–31
- (3) Commission Decision 89/610/EC of 14 November 1989
 (1989) Off. J. Eur. Commun. 351, 39–50
- (4) Commission Decision 93/256/EC of 14 April 1993 (1993)
 Off. J. Eur. Commun. 118, 64–74
- (5) Commission Decision 93/257/EC of 15 April 1993 (1993) Off. J. Eur. Commun. 118, 75–79
- (6) Commission Decision 2002/657/EC of 12 August 2002
 (2002) Off. J. Eur. Commun. 221, 8–36
- (7) André, F., De Wasch, K., De Brabander, H.F., Impens, S., Stolker, L., van Ginkel, L., Stephany, R.W., Schilt, R., Courtheyn, D., Bonnaire, Y., Furst, P., Gowik, P., Kennedy, G., Kuhn, T., Moretain, J.-P., & Sauer M. (2001) *Trends Anal. Chem.* **20**, 435–445
- (8) Commission Decision 87/410/EC of 14 July 1987 (1987) Off. J. Eur. Commun. 223, 18–36
- (9) Guidelines for the Implementation of Decision 2002/657/EC (2008) Doc. SANCO/2004/2726-rev 4, http://ec.europa.eu/ food/food/chemicalsafety/residues/cons_2004-2726rev4_ en.pdf
- (10) Commission Decision 2005/34/EC of 11 January 2005
 (2005) Off. J. Eur. Commun. 16, 61–63
- (11) Council Directive 470/2009/EC of 6 May 2009 (2009) Off. J. Eur. Commun. 152, 11–22
- (12) Stolker, A.A.M., Zuidema, T., & Nielen, M.W.F. (2007) *Trends Anal. Chem.* 26, 967–979
- (13) De Brabander, H.F., Noppe, H., Verheyden, K., Vanden Bussche, J., Wille, K., Okerman, L., Vanhaecke, L., Reybroeck, W., Ooghe, S., & Croubels, S. (2009) *J. Chromatogr. A* **1216**, 7964–7976
- (14) De Brabander, H.F., Le Bizec, B., Pinel, G., Antignac, J.-P., Verheyden, K., Mortier, V., Courtheyn, D., & Noppe, H. (2007) J. Mass Spectrom. 42, 983–998
- (15) Vanhaecke, L., Verheyden, K., Vanden Bussche, J., Schoutsen, F., & De Brabander, H.F. (2009) *LC-GC Europe*, 22, 364–374
- Schürmann, A., Dvorak, V., Crüzer, C., Butcher, P., & Kaufmann, A. (2009) *Rapid Commun. Mass Spectrom.* 23, 1196–1200
- (17) Kaufmann, A., Butcher, P., Maden, K., & Widmer, M. (2008) *J. Chromatogr. A* **1194**, 66–79
- (18) Stolker, A.A.M., Rutgers, P., Oosterink, E., Lasaroms, J.J.P., Peters, R.J.B., van Rhijn, J.A., & Nielen, M.W.F. (2008) *Anal. Bioanal. Chem.* **391**, 2309–2322
- (19) Kolmonen, M., Leinonen, A., Pelander, A., & Ojanperä, I.(2007) Anal. Chim. Acta 585, 94–102
- (20) Fernandez-Alba, A.R., & Garcia-Reyes, J.F. (2008) *Trends Anal. Chem.* 27, 973–990
- (21) Nielen, M.W.F., van Engelen, M.C., Zuiderent, R., & Ramaker, R. (2007) Anal. Chim. Acta 586, 122–129
- (22) Commission Decision 2003/181/EC of 15 March 2003
 (2003) Off. J. Eur. Commun. 71, 17–18
- (23) Commission Decision 2004/25/EC of 22 December 2003
 (2003) Off. J. Eur. Commun. 6, 38–39

- (24) CRL Guidance Paper of 7 December 2007 (2007) CRLs View on State of the Art Analytical Methods for National Residue Control Plans, Community Reference Laboratories (CRLs) for Residues According to Council Directive 96/23/EC, http://www.rivm.nl.bibliotheek/digitaaldepot/ criguidance2007.pdf
- (25) Mooney, M.H., Le Bizec, B., & Elliott, C.T. (2009) Trends Anal. Chem. 28, 665–675
- (26) Scarth, J., Akre, C., Van Ginkel, L., Le Bizec, B., De Brabander, H., Korth, W., Point, J., Teale, P., & Kay, J. (2009) Food Addit. Contam. A 26, 640–671
- (27) Rambaud, L., Bichon, E., Cesbron, N., André, F., & Le Bizec, B. (2005) *Anal. Chim. Acta* 532, 165–176
- (28) Rambaud, L., Monteau, F., Deceuninck, Y., Bichon, E., André, F., & Le Bizec, B. (2007) *Anal. Chim. Acta* 586, 93–104
- (29) Duffy, E., Rambaud, L., Le Bizec, B., & O'Keeffe, M. (2009) Anal. Chim. Acta 637, 165–172
- (30) Maume, D., Deceuninck, Y., Pouponneau, K., Le Bizec, B., André, F., & Paris, A (2001) *APMIS* 109, 32–38
- (31) Maume, D., Le Bizec, B., Pouponneau, K., Solère, V., Paris, A., & André, F. (2003) Anal. Chim. Acta 483, 289–297
- (32) Le Bizec, B., Courant, F., Gaudin, I., Bichon, E., Destrez, B., Schilt, R., Draisci, R., Monteau, F., & André, F. (2006) *Steroids* **71**, 1078–1087
- (33) Antignac, J.P., Le Bizec, B., Monteau, F., & André, F. (2002) Steroids 67, 873–882
- (34) Le Bizec, B., Pinel, G., & Antignac, J.P. (2009) J. Chromatogr. A 1216, 8016–8034
- (35) Noppe, H., Le Bizec, B., Verheyden, K., & De Brabander, H.(2008) Anal. Chim. Acta 611, 1–16
- (36) Ferchaud, V., Le Bizec, B., Monteau, F., & André, F. (1998) *Analyst* **123**, 2617–2620
- (37) Ferchaud, V., Le Bizec, B., Monteau, F., & André, F. (2000) *Rapid Commun. Mass Spectrom.* 14, 652–656
- Hebestreit, M., Flenker, U., Buisson, C., André, F., Le Bizec,
 B., Fry, H., Lang, M., Weigert, A.P., Heinrich, K., Hird, S.,
 & Schanzer, W. (2006) J. Agric. Food Chem. 54, 2850–2858
- (39) Buisson, C., Hebestreit, M., Weigert, A.P., Heinrich, K., Fry, H., Flenker, U., Banneke, S., Prevost, S., André, F., Schaenzer, W., Houghton, E., & Le Bizec, B. (2005) *J. Chromatogr. A* 1093, 69–80
- (40) Bichon, E., Kieken, F., Cesbron, N., Monteau, F., Prevost, S., André, F., & Le Bizec, B. (2007) *Rapid Commun. Mass Spectrom.* 21, 2613–2620
- (41) Becchi, M., Aguilera, R., Farizon, Y., Flament, M.-M., Casabianca, H., & James, P. (1994) *Rapid Commun. Mass Spectrom.* 8, 304–308
- (42) Aguilera, R., Becchi, M., Casabianca, H., Hatton, C.K., Catlin, D.H., Starcevic, B., & Pope, H.G. Jr (1996) *J. Mass Spectrom.* **31**, 169–176
- (43) Cawley, A.T., & Flenker, U. (2008) J. Mass Spectrom. 43, 854–864
- (44) World Anti-Doping Agency, WADA Technical Document-TD2004 EAAS (2004) http://www.wada-ama.org/ rtecontent/document/end_steroids_aug_04.pdf
- (45) International Organization for Standardization (1994) ISO 5725, Accuracy (Trueness and Precision) of Measurement Methods and Results – Parts 1 to 6, http://www.iso.org/iso/ iso_catalogue

- (46) AOAC/FAO/IAEA/IUPAC Expert Consultation (1999) Report on Single-Laboratory Validation of Analytical Methods for Trace Level Concentrations of Organic Chemicals, Miscolc, Hungary, CX/PR 00/11, CRD1 (report of the ad-hoc working group on methods of analysis and sampling)
- (47) Method Validation and Quality Control Procedures for Pesticide Residues Analysis in Food and Feed (2008) Doc. SANCO 3131/2008, http://ec.europa.eu/food/plant/ protection/resources/qualcontrol_en.pdf
- (48) ALINORM 08/31/31 (September 2007) http://www. codexalimentarius.net/download/report/685/al31 31e.pdf
- (49) Report of the Twenty-Third Session of the Codex Committee on Methods of Analysis and Sampling, Budapest, Hungary, 26 February–2 March 2001, http://www.fao.org/ docrep/meeting/005/x9945e/x9945e00.HTM
- (50) Principles and Practices of Method Validation (2000) A. Fajgelj & A. Ambrus (Eds), Royal Society of Chemistry, Cambridge, UK
- (51) Thompson, M., Ellison, S.L.R., & Wood, R. (2002) Pure Appl. Chem. 74, 835–855
- (52) NMKL Procedure 20 (2007) Nordic Committee on Food Analysis, Evaluation of Results from Qualitative Methods, NMKL Secretary General, Oslo, Norway, p. 12, http://www.nmkl.org/Engelsk/Newsletter/70-eng.pdf
- (53) AOAC INTERNATIONAL (1995) J. AOAC Int. 78, 143A
- (54) International Organization for Standardization (2000) ISO/ FDIS 16140:2000, Microbiology of Food and Animal Feeding Stuffs – Protocol for the Validation of Alternative Methods, http://www.iso.org/iso/iso_catalogue
- (55) Jülicher, B., Gowik, P., & Uhlig, S. (1998) Analyst 123, 173–179
- Jülicher, B., Gowik, P., & Uhlig, S. (1999) *Analyst* 124, 537–545
- (57) Gowik, P. (2009) J. Chromatogr. A 1216, 8051-8058
- (58) EURACHEM/CITAC Guide (2007) S.L.R. Ellison & A. Williams (Eds), http://www.eurachem.org
- (59) ISO 11843-2:2000, Capability of Detection—Part 2: Methodology in the Linear Calibration Case, http://www.iso.org/iso/iso catalogue

- (60) Nicholson, J., Lindon, J., & Holmes, E. (1999) *Xenobiotica* 29, 1181–1189
- (61) Want, E., Nordström, A., Morita, H., & Siuzdak, G. (2007) J. Proteome Res. 6, 459–468
- (62) Dumas, M., Debrauwer, L., Beyet, L., Lesage, D., André, F., Paris, A., & Tabet, J.-C. (2002) *Anal. Chem.* 74, 5393–5404
- (63) Dunn, W., Broadhurst, D., Brown, M., Baker, P., Redman, C., Kenny, L., & Kell, D. (2008) *J. Chromatogr. B* 871, 288–298
- (64) Courant, F., Pinel, G., Bichon, E., Monteau, F., Antignac, J., & Le Bizec, B. (2009) *Analyst* 134, 1637–1646
- (65) Kieken, F., Pinel, G., Antignac, J.P., Courant, F., Monteau,
 F., Popot, M.A., Bonnaire, Y., & Le Bizec, B. (2009) *Anal. Bioanal. Chem.* 394, 2119–2128
- (66) Nielen, M., Elliot, C., Boyd, S., Courtheyn, D., Essers, M., Hooijerink, H., van Bennekom, E., & Fuchs, R. (2003) *Rapid Commun. Mass Spectrom.* 17, 1633–1641
- (67) Rijk, J.C.W., Lommen, A., Essers, M.L., Groot, M.J., Van Hende, J.M., Doeswijk, T.G., & Nielen M.W.F. (2009) *Anal. Chem.* 81, 6879–6888
- (68) Lenz, E., & Wilson, I. (2007) J. Proteome Res. 6, 443-458
- (69) Stein, S. (1999) J. Am. Soc. Mass Spectrom. 10, 770-781
- (70) Idborg, H., Zamani, L., Edlund, P.-O., Schuppe-Koistinen, I., & Jacobsson, S. (2005) *J. Chromatogr. B* 828, 14–20
- Jansen, J., Hoefsloot, H., van der Greef, J., Timmerman, M., & Smilde, A. (2005) *Anal. Chim. Acta* 530, 173–183
- (72) Katajamaa, M., & Oresic, M. (2005) Bioinformatics 6, 1-12
- (73) Sysi-Aho, M., Katajamaa, M., Yetukuri, L., & Oresic, M.(2007) *Bioinformatics* 8, 1–17
- (74) Smith, C., Want, E., O'Maille, G., Abagyan, R., & Siuzdak, G. (2006) *Anal. Chem.* 78, 779–787
- (75) Nordström, A., O'Maille, G., Qin, C., & Siuzdak, G. (2006) Anal. Chem. 78, 3289–3295
- (76) Magnus Aberg, K., Torgrip, R., Kolmert, J., Schuppe-Koistinen, I., & Lindberg, J. (2007) J. Chromatogr. A 1192, 139–146
- (77) Lange, E., Tautenhahn, S., Neumann, S., & Gröpl, C. (2008) *Bioinformatics* 9, 375–386

Copyright of Journal of AOAC International is the property of AOAC International and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.