Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 856 (2007) 178–189
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Matrix-comprehensive in-house validation and robustness check of a confirmatory method for the determination of four nitrofuran metabolites in poultry muscle and shrimp by LC–MS/MS

Claudia Bock, Petra Gowik *, Carolin Stachel

Federal Office of Consumer Protection and Food Safety (BVL), EU- and National Reference Laboratory, Mauerstraße 39-42, 10117 Berlin, Germany

Received 3 May 2006; accepted 24 May 2007
Available online 6 June 2007

Abstract

An already well-described method for the determination of nitrofuran metabolites 3-amino-5-methyl-morpholino-2-oxazolidinone (AMOZ), 3-amino-2-oxazolidinone (AOZ), semicarbazide (SEM) and 1-aminohydantoin (AHD) was adapted to the needs of our laboratory and checked for its robustness regarding sample conditions and the processing step. Using the same data, the method was validated and the measurement uncertainty was estimated. All criteria and requirements of Commission Decision 2002/657/EC were fulfilled. The CC_α determined lies between 0.1 and 0.7 μg/kg, the CC_β lies between 0.1 and 0.9 μg/kg, the measurement uncertainty was estimated as being between 7 and 17% taking into account matrix, time and sample preparation influences.

Keywords: Nitrofurans; Metabolite; Validation; Matrix-comprehensive in-house validation; Muscle; Shrimps; Poultry; LC–MS/MS; Residues; Measurement uncertainty; Robustness

1. Introduction

The European Union banned the use of nitrofurans as veterinary drugs for food-producing animals. The minimum required performance limit (MRPL) for nitrofurans is set at 1 μg/kg by Commission Decision 2003/181/EC [1] amending Decision 2002/657/EC [2].

Nitrofurans are resorbed, metabolised and distributed very rapidly. Even shortly after the uptake of nitrofurans only their metabolites are detectable as tissue-bound residues [3–7]. These metabolites act as marker residues for the detection of an illegal use of nitrofurans.

The method, adopted from the method developed by the State Institute for Quality Control of Agricultural Products (RIKILT, Wageningen, NL) [8], describes the determination of all hydrolysable (bound and non-bound) metabolites with an intact side chain. The marker residues 3-amino-2-oxazolidinone (AOZ) for furazolidone, 3-amino-5-methyl-morpholino-2-oxazolidinone (AMOZ) for furaltadone, semicarbazide (SEM) for nitrofurazone and 1-aminohydantoin (AHD) for nitrofurantoin are quantified as their nitrophenyl (NP-) derivatives using the isotopically labelled analogues [1,9,10]. The molecular structures of the parent nitrofurans, their marker metabolites and the corresponding nitrophenyls derivatives are shown in [11] and the principle according to which the nitrofurans react, is shown in [8].

The method presented was modified according to the requirements of the lab also taking into account the equipment and was subsequently validated by matrix-comprehensive in-house validation with an alternative validation concept using the software InterVal. This software uses a fractional factorial design based on an orthogonal experimental plan as described in [12].

By means of the InterVal concept, different factors depending on the samples (e.g. *species*, *condition*/lyophilisation, *processing*) and method (e.g. *operator*, *duration* of sample preparation, *storage* of extracts before analysis) may be determined simultaneously in one validation study and, at the same time, their influence on the measurement results may be assessed. This kind of validation gives details on factors that, if applicable, have to be fixed within the method description or, it shows

* Corresponding author. Tel.: +49 1888 412 2305; fax: +49 1888 412 2300.
E-mail address: petra.gowik@bvl.bund.de (P. Gowik).

1570-0232/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2007.05.044

which kind of samples can be used in which sample condition (fresh/lyophilised, processed/unprocessed). This is necessary for the determination of the fitness-for-purpose. Compared to conventional validation studies, where information about robustness is only available with additional experiments, e.g. during method development, the procedure described provides the possibility to determine robustness within the InterVal validation study.

InterVal is databank-oriented and allows the simultaneous calculation of validation parameters required according to Commission Decision 2002/657/EC [2] like decision limit CC_α and detection capability CC_β, precision, recovery, calibration curve(s) and the respective prediction intervals. Moreover, InterVal delivers the power curve of the method, its measurement uncertainty contributions and a quantitative assessment of the influence of individual factors with regard to dispersion and calibration function [13]. Despite this comprehensive information, compared to conventional validation concepts, e.g. Mottier et al. [14] and national and international standards as e.g. DIN [15,16] and ISO [17] in which, e.g. only one species or matrix is validated without taking other influence factors into account, the number of experiments that has to be performed is smaller. Thus, InterVal enables validation with a small number of samples (a minimum of eight) allowing the finalisation of a study within one or two weeks if needed. On the other hand, the validation study can be interrupted at any time according to the laboratory's necessities and be finished at a later point of time.

Validations of methods for the determination of nitrofuran metabolites in muscle using the alternative InterVal validation concept have not yet been described in publications. Neither is information on the robustness of the method available. So far there are only some publications on validations according to the conventional concept.

Mottier et al. [14] describe, e.g. the validation of a method for the determination of nitrofurans in chicken meat, during which the deuterated internal standards are not added until after the derivatisation with NBA in the form of d_4-nitrophenyl derivatives. Instead of the liquid/liquid extraction we use, a clean-up with SPE is carried out. Within validation, Mottier et al. [14] determine, amongst others CC_α und CC_β, recovery, within- and between-day precision, within-laboratory precision and repeatability.

Finzi et al. [18] use only the two internal standards d_4-AOZ and d_5-AMOZ in their validation for the determination of four nitrofurans in poultry muscle. They specify only some of the validation data required, e.g. accuracy, recovery and, instead of CC_α and CC_β, they indicate the limit of quantification (LOQ).

Furthermore, there are additional method descriptions but without validation with a SPE clean-up e.g. in different species, not defined in detail by Edder et al. [19] as well as in pig muscle and liver by Leitner et al. [20].

In the following sections it is demonstrated which factors have an influence on the measurement results and to what extend they are relevant for the performance parameters of this method. In this context, design factors are analysed as well as noise factors. Additionally, detailed validation results of the method are presented and discussed.

2. Definitions

2.1. Run, InterVal series

A "run" (InterVal series) represents samples of the same individual factor-level combination (Section 2.2), worked up simultaneously with the same sample preparation procedure. This means that one run includes samples of one matrix for the validation calibration curve (Section 2.5), and samples of another matrix of the same kind for the matrix calibration curve (Section 2.6) and the matrix blank sample (Q3) for control purposes as well as one sample without matrix as reagent blank (Q2).

2.2. Factor-level combination

Combination of the factor levels to be applied to one sample.

2.3. Sample preparation procedure

Work-up procedure for all samples (validation calibration and matrix calibration samples, Q2, Q3) of one run with identical factor-level combination.

2.4. Matrix samples

Matrix samples include validation calibration samples as well as matrix calibration samples and matrix blank samples (Q3). The matrix for the validation calibration samples stems from the same species and shows the same condition (lyophilised or fresh) as the validation calibration samples, but it is not identical (see Sections 2.6 and 3.5.2).

2.5. Validation calibration samples, validation calibration

After calculation with an external program and quantification with matrix calibration samples, the validation calibration samples (e.g. S01–S24, each with five fortification levels) constitute a validation calibration curve, which is used by InterVal for the calculation of the required parameters.

2.6. Matrix calibration samples, matrix calibration

An individual matrix calibration curve has to be generated for each series of one factor-level combination (run). The matrix calibration samples are worked up just like the validation calibration samples (Section 2.5). Pre-tests showed that the material used for matrix calibration samples has to originate from the same species and show the same condition (lyophilised or not) as the real samples that have to be quantified; in this case the validation calibration samples. But the material is not identical but most similar. Thus, this procedure of quantification of real samples is applied analogously in the validation study.

3. Experimental

3.1. Chemicals and reagents

Unless indicated otherwise, analytically pure substances and HPLC-grade solvents were used. The standards AHD hydrochloride (99.6%) [2827-56-7], and SEM hydrochloride (≥99%) [563-4-7] were purchased from Sigma–Aldrich (Deisenhofen, Germany) and AOZ (98.3%) [80-65-9] was provided by Riedel-de-Haen (Seelze, Germany). The nitrofuran metabolite AMOZ [43056-63-9] and the labelled metabolites d_4-AOZ hydrochloride (d_4-AOZ), d_5-AMOZ, 1,2-N^{15},C^{13}-SEM hydrochloride (1,2-N^{15},C^{13}-SEM), and $(C^{13})_3$-AHD used as internal standards (>99%) were obtained from Witega (Berlin, Germany). Hexane, methanol and water were, used in HPLC quality, obtained from Fisher Scientific (Wiesbaden, Germany). Ethyl acetate (Suprasolv), 2-nitrobenzaldehyde (2-NBA, p.a.), formic acid (1 M, ultra quality) and ammonium formate (>99%) were purchased from Fluka (Taufkirchen, Germany). Hydrochloric acid (1 M HCl, titrisol quality), sodium hydroxide (2 M NaOH, p.a.) and trisodium-phosphate-10-hydrate (Na_3PO_4, p.a.) were supplied by Merck (Darmstadt, Germany).

The ammonium formate solution (10 mM) used for the mobile phases was adjusted to pH 3.5 with formic acid.

3.2. Standard solutions

Individual standard stock solutions (1 mg/mL) of AHD, SEM, AOZ, and AMOZ, and the labelled internal standards d_4-AOZ, d_5-AMOZ, 1,2-N^{15},C^{13}-SEM, and $(C^{13})_3$-AHD were prepared in methanol. Working solutions (10 ng/mL or 1 ng/mL) were prepared by diluting a stock solution with methanol.

Mixtures of internal standards *IS-Mix* (10 ng/mL) and *NF-Mix 1* with AOZ and AMOZ (10 ng/mL) and *NF-Mix 2* containing SEM and AHD (20 ng/mL) were prepared to be used for sample spiking during method validation.

3.3. Meat samples

Muscle samples from chicken, turkey and shrimp were obtained by several supermarkets and of different producers. The meat was either unprocessed or processed as, for example, marinated or breaded (factor *processing*). As regards processed samples, crumbing or marinade was removed mechanically. The samples were stored at approximately −20 °C until analysis. After thawing, the samples were minced to homogeneity.

Aliquots of the frozen material were lyophilised for at least 72–95 h until constant weight (drying loss between 73.0 and 89.0%) was achieved. The material was then stored in a freezer at −20 °C until analysis.

3.4. Analytical procedure

3.4.1. Sample preparation

The sample preparation procedure used was developed by RIKILT [8] and adapted by the CRL/NRL in Berlin and it is described briefly by Wüst et al. [11]. Therefore, the following description focuses solely on the modifications and information concerning the factor levels chosen.

An amount of 1.00 ± 0.05 g of the homogenised fresh samples or a corresponding amount of lyophilised samples was used (factor *condition*). The lyophilised samples were reconstituted with H_2O.

The internal standard mix *IS-Mix*, the *NF-Mix 1* and *NF-Mix 2* were added to the validation calibration samples and to the matrix calibration samples to obtain the concentrations mentioned in Section 3.5.2.

For derivatisation, hydrochloric acid (0.2 M, 5 mL) and 2-NBA-solution in methanol (0.1 M, 75 μL) were added. The sample was incubated over night at 37 °C either in a thermostated rocking water bath (H_2O) or in a rotating mixer (Rmix) stored in a drying oven (factor *derivatisation*). After a two-fold liquid–liquid extraction with ethyl acetate (4 mL) at pH 7 ± 0.5, the extract was stored over night (4 °C) or was immediately evaporated to dryness (factor *duration*) with a Barkey evaporator or a TurboVap device (factor *evaporation*). The dry residue was redissolved in *n*-hexane (2 mL) and methanol/10 mM ammonium formate, pH 3.5 (15 + 85, v/v, 250 μL). After a two-fold cleaning with *n*-hexane, the resulting extract was filtered and collected in an HPLC vial. The analysis of the samples was carried out with LC–MS/MS either immediately or after storage for 2–3 days at 4 °C (factor *storage*).

3.4.2. LC–ESI-MS/MS

3.4.2.1. HPLC conditions. A binary solvent delivery system (Agilent 1100 Series, Agilent Technologies, Waldbronn, Germany), including a binary pump and a degasser, was used. The liquid chromatograph was equipped with a high-pressure mixing chamber (for the sufficient and reproducible mixing of solvent proportions). Analyses were performed at 30 °C on a Luna C18 column (150 mm × 2.0 mm, 3 μm; Phenomenex, Aschaffenburg, Germany) connected with a Phenomenex C18 pre-column (4 mm × 2 mm). The autosampler was kept at 10 °C and the injection volume was 10 μL. The mobile phase A consisted of methanol and ammonium formate (10 mM, pH 3.5) (v/v, 9 + 1) and the mobile phase B of methanol and ammonium formate (10 mM, pH 3.5) (v/v, 1 + 9). The linear gradient was applied at a flow rate of 0.2 mL/min starting with 20% of A, held for 1 min, increasing to 55% of A within 2 min, increasing again to 95% of A within 6 min, held for 5 min, and decreasing to 20% of A in 1 min, followed by an equilibration duration of 10 min. The retention times for NP-AOZ, NP-AMOZ, NP-AHD, and NP-SEM obtained under these conditions are listed in Table 1.

3.4.2.2. Tandem mass spectrometry. The triple quadrupole mass spectrometer (API 3000, Applied Biosystems, Foster City, CA, USA), equipped with a direct online inlet system and an ESI interface was operated in the positive mode with a TurboIonSpray voltage of 5500 V. The source temperature was adjusted to 400 °C and the entrance potential to 10.0 V. Nitrogen was used as nebulizer gas (13.0 psi), curtain gas (10.0 psi) and collision gas (9.0 psi).

The detection of the analytes was carried out in the multiple reaction monitoring (MRM) by analysing two transitions with

Table 1
Q1 and Q3 mass transitions of the analytes (DP: declustering potential, FP: focusing potential, CE: collision energy, CXP: collision cell exit potential)

Analyte	Retention time (min)	Transition reactions Q1 mass (*m/z*) → Q3 mass (*m/z*)	DP (V)	FP (V)	CE (V)	CXP (V)
NP-d_5-AMOZ	8.32	340.3 → 296.2	46	310	23	8
NP-AMOZ	8.88	335.2 → 291.1	26	200	23	8
		335.2 → 262.1	26	200	25	14
NP-$(C^{13})_3$AHD	10.60	251.9 → 133.9	41	170	19	10
NP-AHD	10.64	249.2 → 134.0	46	290	17	10
		249.2 → 104.1	46	290	33	8
NP-d_4-AOZ	10.62	240.3 → 134.0	46	330	19	10
NP-AOZ	10.70	236.2 → 134.0	46	330	19	10
		236.2 → 104.1	46	330	31	8
NP-$C^{13}(N^{15})_2$-SEM	10.87	212.2 → 168.0	36	250	15	12
NP-SEM	10.89	209.2 → 191.9	36	260	17	16
		209.2 → 166.0	36	260	15	14

a dwell time of 100 ms with the resolution of Q1 and Q3 set to "unit". Transition reactions given in Table 1 as well as the adjustments of the declustering potential (DP), focusing potential (FP), collision energy (CE) and collision cell exit potential (CXP) of the analytes.

3.5. Validation experiment

The validation was performed in accordance with Commission Decision 2002/657/EC, Technical Annex, Chapter 3.1.3. applying the alternative validation approach [2]. Because the MRPL is 1 μg/kg, the concentration levels for the validation calibration samples for all nitrofurans were validated close to zero [2].

3.5.1. Establishment of an experimental design

For the establishment of the experimental design, the commercially available software InterVal (quodata GmbH, Dresden, Germany) was used.

In reality, it is not possible to look at a total population with all its samples and to be able to verify a representative sample. The advantage of the InterVal validation concept and the application of a stratified experimental plan is that by means of the targeted selection of factors, varied on two factor levels (cf. Table 2) which means that with this also the borders are covered, a hypothetic total population is obtained. This hypothetic population covers the most important characteristics of a representative population [12].

Therefore, the selection of factors is exemplary and exclusively dependent on the requirements and characteristics of the method and the purpose for which it is to be used. A method for the analysis of nitrofuran metabolites in turkey and chicken muscle and shrimp was to be validated. The species was defined as leading factor and varied on three levels. For statistical reasons, each level of the leading factor requires eight different samples. Additionally, no more than seven further factors varying on two levels were needed [21]. The limitation to 7 + 1 factors results from the statistical experimental plans used in InterVal. They are based on a fractioned $2^{\{7-4\}}$-plan, which allows the orthogonal variation of seven factors on two factor levels each with only eight factor level combinations [12,21].

The selection of the factors depends not only on the method which is to be validated, but also on the individual conditions in the laboratory in which the method is to be established. The factor levels of the eight factors were defined as follows (see Table 2):

The factor *duration* of the sample preparation (2d/3d) was chosen to simulate a break in sample preparation, a situation that may occur in the everyday laboratory routine. Sample preparation (normally 2d) was interrupted after the ethyl acetate extraction until the next day (3d). The factor *storage* of sample extracts (no/yes) was supposed to simulate a situation when the sample extracts cannot be analysed immediately after sample preparation and have to be stored for 2 or 3 days at +4 °C before analysis. For a detailed description of the other factors and factor levels see Sections 3.3 and 3.4.1.

Table 2
Factor and factor levels

Leading factor:	1	Species	Shrimps, turkey, chicken	
Factor:	2	Operator	A	B
	3	Processing	Processed (proc.)	Unprocessed (unproc.)
	4	Condition	Lyophilised (lyo)	Fresh (fresh)
	5	Derivatisation	Rocking water bath (H_2O)	Rotating mixer (Rmix)
	6	Evaporation	Barkey (Barkey)	TurboVap LV (T-Vap)
	7	Duration	2 days (2d)	With interruption (3d)
	8	Storage	0 days (no)	2–3 days at +4 °C (yes)

Table 3
Experimental design (proc.: processed, unproc.: unprocessed, H_2O: rocking water bath, Rmix: rotating mixer, T-Vap: Turbo-Vap)

Run	Sample name	1 Species	2 Operator	3 Processing	4 Condition	5 Derivatisation	6 Evaporation	7 Duration	8 Storage
S24	P040142	Chicken	A	proc.	Fresh	H_2O	T-Vap	3 d	No
S08	P040144	Shrimps	A	proc.	Fresh	Rmix	T-Vap	3 d	Yes
S17	P040141	Chicken	B	unproc.	Fresh	H_2O	T-Vap	2 d	Yes
S12	P020367	Turkey	B	proc.	Fresh	Rmix	T-Vap	3 d	No
S07	P040306	Shrimps	A	proc.	Fresh	H_2O	Barkey	2 d	No
S23	P040136	Chicken	A	proc.	Fresh	Rmix	Barkey	2 d	Yes
S10	P040139	Turkey	B	unproc.	Lyo	H_2O	T-Vap	3 d	Yes
S18	P030413	Chicken	B	unproc.	Fresh	Rmix	Barkey	3 d	No
S14	P030326	Turkey	A	unproc.	Fresh	H_2O	Barkey	3 d	No
S05	P040173	Shrimps	A	unproc.	Lyo	Rmix	Barkey	2 d	Yes
S11	P040137	Turkey	B	proc.	Fresh	H_2O	Barkey	2 d	Yes
S01	P010509	Shrimps	B	unproc.	Fresh	Rmix	T-Vap	2 d	No
S21	P040138	Chicken	A	unproc.	Lyo	H_2O	Barkey	2 d	No
S16	P040111	Turkey	A	proc.	Lyo	Rmix	Barkey	3 d	Yes
S02	P030414	Shrimps	B	unproc.	Fresh	H_2O	Barkey	3 d	Yes
S19	P040282	Chicken	B	proc.	Lyo	Rmix	T-Vap	2 d	No
S15	P020400	Turkey	A	proc.	Lyo	H_2O	T-Vap	2 d	No
S22	P030541	Chicken	A	unproc.	Lyo	Rmix	T-Vap	3 d	Yes
S06	P040472	Shrimps	A	unproc.	Lyo	H_2O	T-Vap	3 d	No
S13	P030494	Turkey	A	unproc.	Fresh	Rmix	T-Vap	2 d	Yes
S03	P040136	Shrimps	B	proc.	Lyo	H_2O	T-Vap	2 d	Yes
S09	P040121	Turkey	B	unproc.	Lyo	Rmix	Barkey	2 d	No
S20	P040299	Chicken	B	proc.	Lyo	H_2O	Barkey	3 d	Yes
S04	P020399	Shrimps	B	proc.	Lyo	Rmix	Barkey	3 d	No

From this data, InterVal creates an experimental design, which shows 24 different runs (S01–S24), each with an individual sample preparation procedure based on one of the defined factor-level combinations (Table 3). To be able to realise these experiments during laboratory routine, the experimental order suggested by InterVal produced through a randomisation process was slightly changed.

3.5.2. Samples and calibration

Each of the 24 runs (S01-S24) consisted of one sample matrix, which was divided into the validation calibration samples (Section 2.5) spiked at five fortification levels. Quantification was performed using individual matrix calibration curves (Section 2.6) for each run. Therefore, in addition to the five validation calibration samples, another seven matrix calibration samples (Section 2.4) and one matrix blank (Q3) were required. Additionally, a reagent blank (Q2) was needed; thus, a total of 14 samples were worked up for one run.

The calibration curves of the validation calibration samples (Section 2.5) for AOZ and AMOZ or SEM and AHD were produced in a concentration range of 0.1–0.4 μg/kg or 0.5–2.0 μg/kg. The matrix calibration curves for AOZ and AMOZ or SEM and AHD were produced in a concentration range of 0.1–0.6 μg/kg or 0.4–2.4 μg/kg.

4. Results/discussion

InterVal validation provides parameters like the critical concentrations CC_α and CC_β, repeatability s_r, within-laboratory reproducibility s_{wR}, recovery, calibration curves, and prediction interval and additionally optional parameters like measurement uncertainty, uncertainties of individual factors, power function, factorial effects and boxplots. With this comprehensive information on robustness and the scope of the method is provided. For the calculation of these parameters, InterVal uses the analyte contents of the validation calibration samples, calculated by matrix calibration.

The other parameters required by Commission Decision 2002/657/EC [2] such as selectivity/specificity were already obtained during method adaption as well as first information on robustness and scope of application. Information on stability is obtained in additional experiments. Since trueness (Section 4.2) can only be determined with the help of certified reference material, which was not available in this case, the recovery (Section 4.7) was determined as corrected recovery by the use of internal standards added to the blank matrix samples.

4.1. Selectivity/specificity

The specificity of the method was demonstrated by analysing structurally related compounds like the labelled internal standards (d_4-AOZ, d_5-AMOZ, 1,2-N^{15},C^{13}-SEM and $(C^{13})_3$-AHD) which can be separated from the analyte by chromatography and spectroscopy (cf. Fig. 1). The influence of matrix interference was investigated by analysing the 24 blank matrix samples different in type and condition. The same material was used for the corresponding validation calibration samples with the corresponding factor level combination (Table 3).

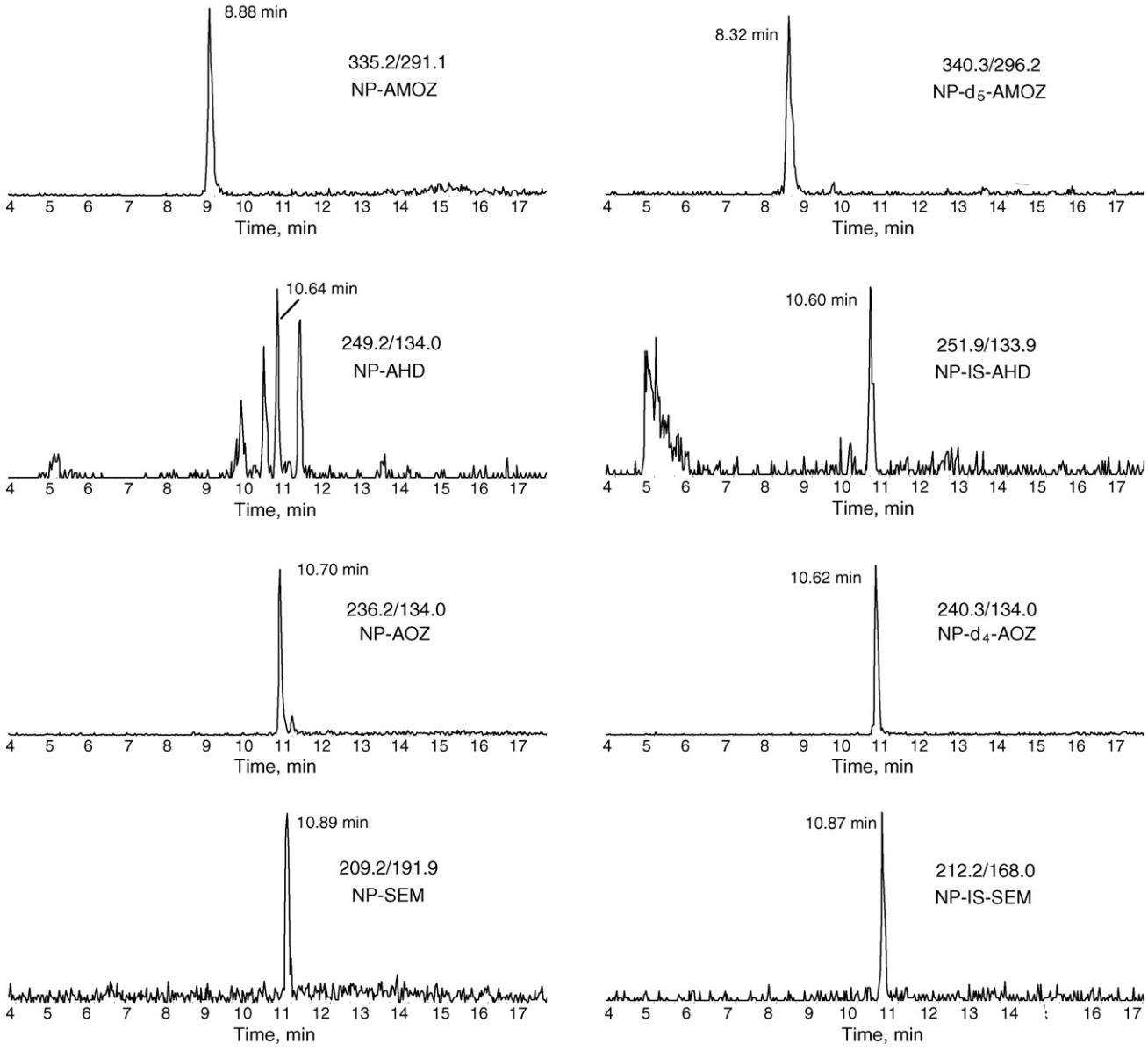

Fig. 1. Ion traces of the NP-derivatives in turkey (NP-AOZ and NP-AMOZ: 0.1 μg/kg; NP-AHD and NP-SEM: 0.5 μg/kg; internal standards: 0.5 μg/kg), chromatographic conditions under Section 3.4.2.

4.2. Trueness and robustness/scope of application

Instead of trueness, the recovery (Section 4.7), also referred to as corrected recovery, was determined by means of spiked blank matrix and the use of internal standards.

The applicability and robustness of the method were proven by analysing eight samples of different origins and of three species (factor *species*: shrimp, turkey and chicken). The material may be processed (marinated or with crumbing) or unprocessed (factor *processing*) and, at the same time, lyophilised or fresh (factor *condition*).

The method proved to be robust concerning the factors *operator*, *derivatisation*, *evaporation*, *storage* and *duration*.

The validation study was performed over a period of two months, hence the analytical system can be considered stable for this time period at least. The method is applicable to shrimps, turkey and chicken. For confirmatory analysis, matrix calibration with the same species and the same matrix condition (lyophilised or fresh) is required due to different matrix suppression effects and recovery. In additional experiments, it could be demonstrated that the recovery of the internal standards also depends on the matrix. Quantitation with a standard calibration is, e.g. not possible for AHD due to reduced recovery of the $(C^{13})_3$-AHD in validation calibration samples, which results in a higher recovery of AHD. Due to different recoveries of d_4-AOZ in lyophilised compared to fresh turkey muscle, a matching, lyophilised or fresh, matrix calibration curve is obligatory.

In conclusion, to prevent enhanced or reduced recovery, the generation of a matrix calibration curve from the same matrix as the samples that have to be quantified and the use of internal standards are mandatory for quantitative purposes.

In general, the slopes of the 24 validation calibration curves depend on the matrix and the detected analyte. The slopes are between 0.92 and 1.09.

4.3. Stability

At −25 °C the analytes in matrix are stable for at least 12 months. This was tested with incurred material which was used for the proficiency test NIFU_0603 organised by the CRL/NRL Berlin, Germany in 2003. In standard solution, the analytes are stable for at least 12 months when stored at +4 °C in the dark. This validation showed that the derivatised extracts of the samples, kept in properly sealed HPLC-vials in the refrigerator at +4 °C are stable for a minimum of 3–4 days (comment, see Section 4.5).

4.4. Outlier tests

InterVal allows the application of the Grubbs-test for individual measurement values as well as for calibration functions, the application of the Cochran-test to check the scattering in the calibration curves, and finally the performance of tests to register factorial effects [13].

For AHD, an outlier was recognised on the third fortification level (1.5 μg/kg). Because this outlier could not be explained by experiment, it was not excluded from the evaluation.

4.5. Confirmation

The lowest fortification level of the validation calibration was chosen in such a way that at least 50% of the samples are confirmable. For SEM and AHD, this concentration was 0.5 μg/kg. For AOZ and AMOZ, for the detection of which the method was found to be more sensitive, it was 0.1 μg/kg.

All samples were confirmable in the case of AOZ. AMOZ could be confirmed in all samples, except in one at 0.1 μg/kg. For SEM and AHD, only two samples each could not be confirmed (SEM: 0.5 and 1.0 μg/kg; AHD: 0.5 and 2.0 μg/kg). Only samples of three different matrices were affected by the lack of confirmability: unprocessed, lyophilised turkey (S10), unprocessed, lyophilised chicken (S22) and processed fresh chicken (S23), the sample extracts of which were stored for several days before the analysis. Nevertheless all values, including the ones which could not be confirmed, were taken into account in the evaluation. This is permissible because only one sample each, unconfirmable for SEM (1.0 μg/kg) or AHD (2.0 μg/kg), lies above the corresponding CC_β (0.82 and 0.88 μg/kg). Thus, the share of unconfirmed samples above the CC_β is about 1% (one of 96 validation calibration samples) and lies below the maximal permitted β-error of 5% [2].

4.6. Critical concentrations (CC_α, CC_β)

The critical concentrations for the analytes were calculated after recovery correction by means of matrix calibration curves and by using internal standards calculated by the worst case scenario. This means that an extrapolation at the lowest spiked concentration level is carried out by means of parallel projection unto the *y*-axis [22]. For all analytes, CC_α and CC_β are below the MRPL 1.0 μg/kg (see Table 4).

Table 4
Critical concentrations for the analytes

	AOZ	AMOZ	AHD	SEM
Decision limit (CC_α) (μg/kg)	0.12	0.13	0.67	0.70
Detection capability (CC_β) (μg/kg)	0.14	0.15	0.82	0.88

Mottier et al. [14] calculated CC_α and CC_β following the calibration curve procedure as explained in Commission Decision 2002/657/EC [2] as the conventional validation procedures. In this procedure, blank matrices of chicken meat were spiked at different concentration levels and a linear extrapolation results in the CC_α. Nevertheless, the experiment of Mottier et al. [14] take time effects into account as being performed within a one-month period by different operators and different instrument performances, but it does not consider the matrix mismatch effects [23,24] in any way. In their study, CC_α and CC_β were 0.21 and 0.36 μg/kg, 0.12 and 0.21 μg/kg, 0.11 and 0.19 μg/kg, 0.20 and 0.34 μg/kg for AHD, AMOZ, AOZ and SEM, respectively.

Compared to this one CC_α and CC_β resulting from our study were similar for AOZ and AMOZ, but higher for AHD and SEM. The main reason for a different method sensitivity is certainly the different calculation mode. The alternative validation concept in which three species and seven further factors have been validated at the same time, may result in higher CC_α and CC_β, due to the higher variability of measurement results coming from the choice of factor and from differences in the clean-up procedure (liquid/liquid extraction or SPE). Additionally, the completely different concept and the mathematical–statistical calculations of CC_α as well as CC_β and the application of the worst case scenario produce different values for the validation parameters [22].

In conventional validation, Finzi et al. [18] verified a limit of quantification (LOQ) for poultry matrix for the four nitrofuran metabolites of 0.5 μg/kg with an accuracy between 70 and 130% (CV < 20%). The LOQ was calculated by spiking a matrix blank at 0.3, 0.5, 1.0 and 1.5 μg/kg and was verified by 20 blank samples fortified with AOZ, AMOZ, AHD and SEM to a final concentration of 0.5 μg/kg. The limit of detection (LOD) for AMOZ was around 1 μg/kg and for the other three metabolites close to 0.2 μg/kg [18]. This study is not taking into account any time or matrix effects.

Therefore, LOD and LOQ cannot be compared directly with CC_α and CC_β. In contrast to the results of Finzi et al. [18] our validation results presented show lower values for CC_α and CC_β for AOZ and AMOZ whereas they are higher for AHD and SEM. It has to be pointed out again that by means of the validation of the three species chicken, turkey and shrimps at different stages of processing as, e.g. unprocessed or marinated/breaded, a greater variation of the values can be expected. This also means that the calculated CC_α und CC_β will be higher than in the case of a validation of only one matrix but at the same time, this comes closer to the real conditions. Real samples also vary to a

great extend and do not consist of the same or the same type of material as it is assumed in conventional validation studies.

4.7. Repeatability s_r, within-laboratory reproducibility s_{wR} and recovery

In the form of a table (e.g. Table 5), InterVal indicates the calculated repeatability s_r, the within-laboratory reproducibility s_{wR} and corrected recovery for the fortification levels and, by means of interpolation also for the concentrations between the fortified concentration levels.

Generally, the corrected recovery, lies in the range of 93–101% for all analytes in a concentration range below 1 μg/kg and thus fulfils the requirements of the Commission (−50 to +20%) [2].

The repeatability s_r calculated in InterVal is naturally higher for lower concentrations than for higher ones. In all cases, it lies below 17% for all analytes. The same applies for the within-laboratory reproducibility s_{wR}. For lower concentrations, it lies within the same range as the repeatability s_r. But in the case of higher concentrations, as expected, the within-laboratory reproducibility is higher than the repeatability. At lower concentrations, random deviations, which may e.g. be due to peak integration or sample handling, have a stronger effect on the results than at higher concentrations where matrix effects have a greater impact.

Mottier et al. [14] determined recoveries between 85 and 122%, a within-laboratory precision for AHD, AMOZ and AOZ ≤15% and less satisfactory for SEM (≤35%). The repeatability at the 95% confidence level was calculated from the within-day precision [14]. For the validation for the determination of nitrofurans in poultry muscle, Finzi et al. [18] only use the internal standard d_5-AMOZ for the quantification of AMOZ and d_4-AOZ for the quantification of AOZ, AHD and SEM. They determined a recovery of 30% for NP-AOZ and NP-SEM as well as about 100% for NP-AHD for a single ethyl acetate extraction step. As said above, the results of the recoveries of the validation study presented lie much closer to 100% and with this show that especially in the case of confirmatory methods, the use of matrix calibration is of great importance and strongly advisable despite the greater effort.

Table 5
Relative coefficient variation (CV) of repeatability s_r, within-laboratory reproducibility s_{wR} and corrected recovery of AOZ, AMOZ, AHD and SEM

	CV s_r (%)	CV s_{wR} (%)	Corrected recovery[a] (%)
AOZ concentration level (μg/kg)			
0.100	9.1	9.1	92.6
0.150	7.2	7.8	95.2
0.200	6.1	7.3	96.5
0.250	5.3	6.9	97.3
0.300	4.8	6.6	97.8
0.350	4.4	6.3	98.2
0.400	4.1	6.1	98.5
AMOZ concentration level (μg/kg)			
0.100	13.1	13.1	92.9
0.150	9.0	9.3	96.3
0.200	7.0	8.1	98.0
0.250	5.7	7.5	99.0
0.300	4.9	7.2	99.7
0.350	4.3	7.0	100.2
0.400	3.8	6.9	100.6
AHD concentration level (μg/kg)			
0.500	13.8	13.9	94.9
0.750	11.0	11.0	95.7
1.000	9.4	9.5	96.1
1.250	8.3	8.9	96.4
1.500	7.5	8.6	96.5
1.750	6.9	8.4	96.6
2.000	6.5	8.3	96.7
SEM concentration level (μg/kg)			
0.500	16.7	16.7	97.0
0.750	12.5	12.5	98.4
1.000	10.2	11.4	99.1
1.250	8.8	11.0	99.6
1.500	7.7	10.7	99.8
1.750	6.9	10.5	100.0
2.000	6.3	10.3	100.2

[a] The recovery was corrected by the use of internal standard and matrix calibration.

4.8. Matrix calibration curves

The samples for matrix calibration were fortified in the range given in Section 3.5.2.

4.8.1. Sensitivity

The mean value calculated for the sensitivity b_0 (standardised slope) of the matrix calibration curve multiplied with the internal standard concentration was calculated for AOZ, AMOZ, SEM, AHD, and the corresponding internal standards (IS). Only in the case of AHD, the deviation of the standardised slope was rather high with 16.6%. The deviations of the slopes of all other calibration curves were below 7.7%.

4.8.2. Linearity

The linearity of the individual calibration curves was demonstrated for all of the 24 matrix calibration curves. Correlation *r* was better than 0.99 in most of the cases. Only for SEM (five cases) and AHD (four cases) the correlation *r* was between 0.98 and 0.99.

4.9. Validation calibration curves

The prediction intervals of the validation calibration curves of AOZ and AMOZ are smaller than the ones of AHD and SEM, i.e. the results vary less for AOZ and AMOZ, the method is more robust for these two analytes.

InterVal presents validation calibration curves presented in a discriminable way according to the choice of factor (e.g. Fig. 2). This way, for calibration curves with a significantly deviating slope, it can be determined which sample is affected or which factor has this influence.

In the calibration curves which represent the *species* shrimps, turkey and chicken, no tendency could be observed. Thus, the factor *species* did not have any influence on the results. Neither does the factor *operator* for AOZ, AMOZ and SEM show any

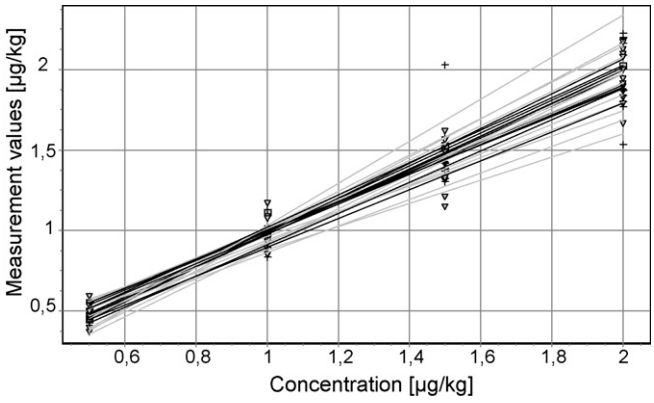

Fig. 2. Sample calibration curves of the factor *operator* A (black) or B (grey) for AHD.

influence. But for AHD (cf. Fig. 2), tendencies can be observed: the results of *operator* A show less variation as compared to those of *operator* B. But as it can be seen later on (Sections 4.10 and 4.13), this effect is not statistically significant.

The factor *processing* shows an effect for SEM which it does not for the other analytes: the results of processed material vary more than the ones of unprocessed matrix. All the other factors *derivatisation, duration*, *condition*, *evaporation*, and *storage* do not show any significant effect for any of the analytes.

4.10. Measurement uncertainties within the validated calibration range

Within a diagram, InterVal gives an estimation of the measurement uncertainty contributions of the individual factors, the within-laboratory reproducibility standard deviation (relative combined uncertainty) and the relative repeatability standard deviation (e.g. Fig. 3). InterVal shows, dependent on the concentration, the measurement uncertainty contributions of the individual factors, which are reflected in the relative matrix/run S.D. and thus also in the within-laboratory reproducibility standard deviation. The latter results from the "sum" of the measurement uncertainties of the individual factors and the time effect.

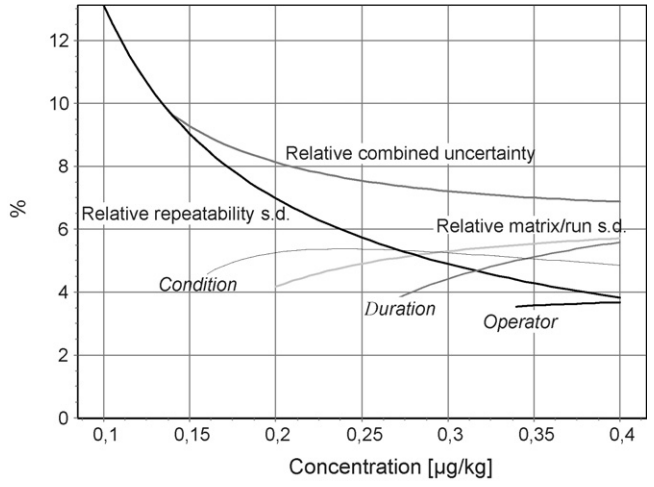

Fig. 3. Measurement uncertainties of AMOZ.

It can be noticed (Fig. 3) that the measurement uncertainty (relative combined uncertainty) and the uncertainty of the repeatability (relative repeatability s.d.) decrease, as the concentration rises and that, for lower concentrations, they lie in the range of 13%, 9%, 14% and 17% for AMOZ, AOZ, AHD, and for SEM, respectively.

For AMOZ and AOZ, the two factors which have a visible influence are the factor *duration* and *condition*. Their measurement uncertainties lie below 6% each. In the case of AMOZ, the factor *operator* additionally seems to have a slight influence on the variation of the measurement results, which was however not detectable in the validation calibration curves (Section 4.9) and which has no statistically significant effect as shown in Table 6.

For AHD, the factors *duration* and *evaporation* also play an important role for the measurement uncertainty but they do also not exceed 6%. The factor *operator* shows a measurement uncertainty of under 3% and therefore it is not indicated in the figure of AHD, although the results showed a smaller variation in the validation calibration curves (Section 4.9) on one level compared to the other one.

The factors *operator* and *processing* contribute to the measurement uncertainty for semicarbazide. But the contributions of the different factors to the measurement uncertainty of the individual analytes are not significant as described in Sections 4.11 and 4.13.

Whether the factors mentioned above, which mainly contribute to measurement uncertainty, have a significant effect was checked separately as described in Sections 4.11, 4.12 and 4.13.

The measurement uncertainty and partially the individual measurement uncertainty components are taken into account only in a few methods in the literature, although it provides important information on the measurement results, especially with regard to the decision whether a limit is exceeded.

Mottier et al. [14] estimated the measurement uncertainty of the results of in-house tests of spiked samples. Every step of the sample preparation was allocated with a defined measurement uncertainty and resulted in an expanded measurement uncertainty calculated with 2 as expansion factor for a 95% confidence interval. For AHD, the expanded measurement uncertainties lie in the range of 22–26% (fortification levels: 0.46–0.93 μg/kg), for AMOZ in the range of 12–21% (fortification levels: 0.60–1.20 μg/kg), for AOZ in the range of 9–32% (fortification levels: 0.43–0.87 μg/kg), and for SEM in the range of 26–36% (fortification levels: 0.36–0.72 μg/kg).

The validation study presented, where lower spike levels were used for AMOZ and AOZ, but higher ones for AHD and SEM, gives lower calculated expanded measurement uncertainties for AOZ (≤12%, fortification levels: 0.1–0.4 μg/kg), similar or lower values for AMOZ (≤14%, fortification levels: 0.1–0.4 μg/kg, e.g. Fig. 3), and comparable values for AHD (≤28%, fortification level: ≤0.5 μg/kg), and for SEM (≤33%, fortification level: ≤0.5 μg/kg) for the concentration ranges covered by the fortification levels of Mottier's and our investigations. In this context, it can be observed that the measurement uncertainties fall to a constant low value (e.g. Fig. 3) as the

Table 6
Matrix-induced deviation of recovery in percentage for AMOZ, AOZ, SEM and AHD

	+	−	AMOZ		AOZ		SEM		AHD	
			proport.	const.	proport.	const.	proport.	const.	proport.	const.
Shrimp			2.638	1.814	0.083	0.594	7.225	4.616	0.215	0.474
Turkey			−4.758	−3.343	−1.094	−1.012	−6.462	−3.670	5.354	0.558
Chicken			2.120	1.529	1.010	0.418	−0.763	−0.946	−5.569	−1.032
Operator	A	B	−5.383	−2.643	−2.402	−2.330	−3.735	−1.511	−1.754	1.497
Processing	unproc.	proc.	−5.836	−1.255	−0.942	−0.974	−8.333	−5.391	−3.092	−0.168
Condition	fresh	lyo	−0.576	2.225	−2.884	−0.815	7.389	0.723	5.948	4.604
Derivatisation	Rmix	H_2O	−4.657	−0.778	1.786	2.901	−3.316	−1.148	−4.004	−2.662
Evaporation	T-Vap	Barkey	3.016	1.613	2.501	−0.883	−3.164	−1.398	−0.490	0.861
Duration	2 d	3 d	6.331	2.653	4.942	4.819	5.467	1.888	−4.752	0.722
Storage	no	yes	0.296	−1.906	−0.863	−2.132	−1.327	−2.353	0.440	−0.414

concentration rises. In Mottier's approach no tendency could be observed due to calculations only at the four fortification levels.

The measurement uncertainty of the standard solutions (Section 3.2) is calculated on the basis of the weighing in of 10 mg, of taking into account the purity of the standards used, of three required dilution steps with methanol, of having regard to the expansion coefficient, of the use of a 10 mL volumetric flask. This calculation results in a measurement uncertainty of 8.6%. This measurement uncertainty has to be added by means of the error propagation law into the calculation of measurement uncertainties by InterVal in the case of the use of different standard solutions. When always the same solution is used, its uncertainty needs not to be included into the total measurement uncertainty.

4.11. Uncertainties of individual factors

Uncertainties of individual factors which influence the results are not available with a conventional validation. But InterVal provides a scheme (see Table 6), which allows to make a statement about the uncertainties of individual factors.

The first three lines show the matrix-induced deviations of recovery of the leading factor *species* for the levels shrimps, turkey and chicken. In general, the values for *proport.* (concerning the slopes of the calibration curves) lying below 10% and the values for *const.* (concerning the position of the calibration curves) below 10% can be neglected. The relative deviations of the slopes (*proport.*) of the three leading factors in this study lie below ±7.2%, with the highest deviation in the case of semicarbazide, indicating that there might be a different behaviour between shrimps (*proport.*: 7.225) and turkey (*proport.*: −6.462). Nevertheless the difference is too small to require the application of different methods. Especially since the precision parameters also lie in the required range and CC_α and CC_β are below the MRPL and therefore comply with legal requirements.

For the other factors, apart from the leading factor *species*, the *proport.* (factorial slope effect) indicates differences between the mean slope of all calibration curves, the slopes of the "+" level variation and the "−" level variation of the respective factor. The same is valid for the *const.* (factorial mean effect), which indicates the difference between the means of the calibration levels of the "+" level variation and the "−" level variation of the respective factor. Factorial slope effects and factorial mean effects below 10% can be neglected.

4.12. Power function

InterVal also calculates the power function, with the slope of the power curve depending on the analyte, the kind of method (screening or confirmation), the limits prescribed and the scatter of measurement results determined. It provides detailed information on the probability of detecting a contaminated sample per concentration. With this InterVal represents an instrument for method assessment and method comparison.

The power curves for AOZ (cf. Fig. 4) and AMOZ are steeper than those for AHD and SEM. This means that CC_α (∼50% probability) and CC_β (95% probability) lie closer together in the case of AOZ and AMOZ and the method works better for these two analytes, i.e. due to the smaller variation range for SEM and AHD, contaminated samples can be detected at lower concentrations.

4.13. Further parameters which can be determined by InterVal

InterVal provides other parameters like *t*-values (factorial effects), boxplots of measurement and boxplots of residuals.

Because the *t*-values lie between −2 and +2, within the validated concentration range, a significant effect on the different species was not produced by any of the factor level of all the factors and for all analytes The boxplots derived from the data of this validation study do not show any tendencies or significant effects.

With this, the robustness of the method could be proven. All factors examined in this validation did not show any significant effects. Therefore it is possible to stop sample preparation after the ethyl acetate extraction and store the sealed extract at 4 °C until the next day. In addition, it is possible to store samples up to 3 days at 4 °C before the analysis, if the samples cannot be analysed directly after the preparation. Nevertheless this should be avoided to prevent difficulties that might occur in the confirmation of samples with low concentrations of SEM and AHD (see Section 4.5).

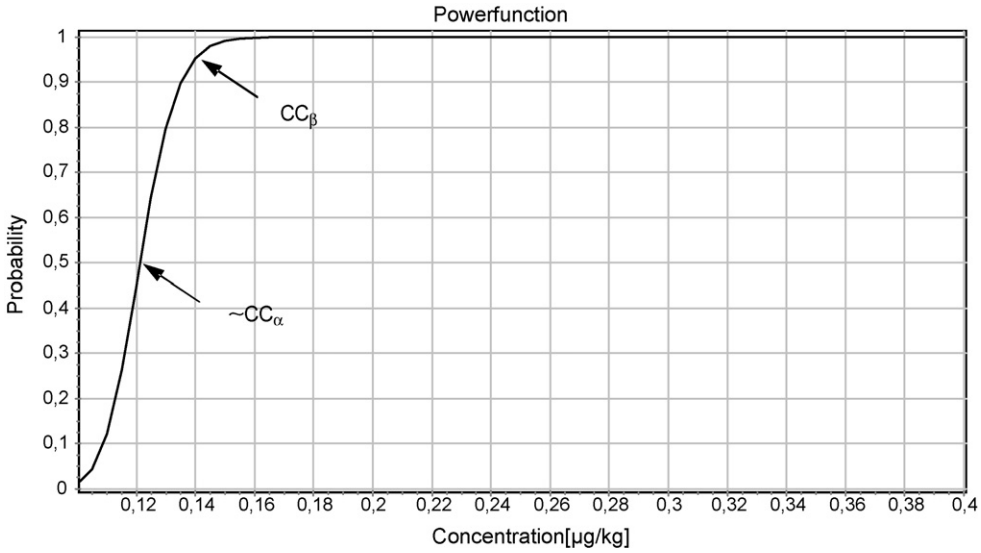

Fig. 4. Power curve of AOZ.

5. Conclusion

The intention of each validation study is the verification of the fitness for purpose and the highest robustness achievable of a test method.

A validation with InterVal allows an efficient, comprehensive and statistically founded statement on the characteristics of the method. The most relevant variations of samples which might occur in daily routine can be simulated as defined factors which can be adapted to the individual needs of each laboratory. Amongst these are, e.g. the influence of different matrix conditions (e.g. lyophilisation, processing) and method parameters on the measurement result, and, with this, statements on the robustness of the method.

In conclusion, after this comprehensive validation and robustness study, the method is checked intensively, very well described concerning probable influences and it can be stated with great certainty that the method is applicable to AOZ, AMOZ, AHD and SEM under varying matrix and environmental conditions. The procedure can safely be used for fresh or lyophilised (factor *condition*) poultry muscle (chicken and turkey), and shrimp regardless whether the products are unprocessed (fresh) or processed like marinated or breaded muscle (factor *processing*).

Matrix calibration, using fresh matrix of identical species has to be applied to avoid enhanced or reduced recovery. For correct quantification lyophilised samples also require matrix calibration using lyophilised material. Therefore, all the tested variations may be introduced safely without influencing the result of the evaluation of a sample (compliant or not compliant).

With regard to the conditions described above it has to be stated that all the determined validation parameters fully comply with the criteria demanded in Commission Decision 2002/657/EC [2].

In addition to that InterVal provides a realistic assessment of the measurement uncertainty and also includes essential components of measurement uncertainty such as matrix, time and influences originating from sample preparation.

Acknowledgements

The authors wish to thank Mrs. A. Neumärker and Mr. A. Petruck for their technical assistance in the investigations.

References

[1] Commission Decision (EC) 2003/181/EC, Off. J. Eur. Comm., No. L71 (2003) 17.
[2] Commission Decision (EC) 2002/657/EC, Off. J. Eur. Comm., No. L221 (2002) 8.
[3] R.J. McCracken, D.G. Kennedy, J. Chromatogr. B 691 (1997) 87.
[4] R.J. McCracken, W. Blanchflower, C. Rowan, M. McCoy, D.G. Kennedy, Analyst 120 (1995) 2347.
[5] J.F. Nouws, J. Laurensen, Vet. Quart. 12 (1) (1990) 56.
[6] K.M. Cooper, P.P.J. Mulder, J.A. van Rhijn, L. Kovacsics, R.J. McCracken, P.B. Young, D.G. Kennedy, Food Addit. Contam. 22 (5) (2005) 406.
[7] L. Hoogenboom, M. van Kammen, M. Berghmans, J. Koeman, H. Kuiper, Food Chem. Toxicol. 29 (5) (1991) 321.
[8] J.A. van Rhijn, P.P.J. Mulder, RIKILT (State Institute of Quality Control of Agricultural Products), Muscle of poultry, rabbit and aquaculture products – detection, identification and quantification of furazolidone, furaltadone, nitrofurantoin and nitrofurazone – LC–MS/MS confirmatory analysis, Meeting of the National Experts on 11 April 2002, Brussels.
[9] E. Horne, A. Cadogan, M. O'Keeffe, L.A.P. Hoogenboom, Analyst 121 (1996) 1463.
[10] R.J. McCracken, D.G. Kennedy, Food Addit. Contam. 14 (1997) 507.
[11] B. Wüst, C. Sauber, H. van Rhijn, Quantitation of nitrofuran metabolites in shrimp and poultry by LC/MS/MS using Agilent LC/MSD Trap XCT, Agilent Application, Food (2004).
[12] B. Jülicher, P. Gowik, S. Uhlig, Analyst 124 (1999) 537.
[13] S. Uhlig, P. Gowik, W. Radeck, Anal. Chim. Acta. 483 (2003) 351.
[14] P. Mottier, S.-P. Khong, E. Gremaud, J. Richoz, T. Delatour, T. Goldmann, P. Guy, J. Chromatogr. A 1067 (2005) 85.

[15] Deutsches Institut für Normung, DIN 55350, Teil 34, Beuth Verlag, Berlin (1991).
[16] Deutsches Institut für Normung, DIN 32645, Beuth Verlag, Berlin (1994).
[17] ISO/DIS 11843-1, International Standards Organisation, Geneva (1995).
[18] J. Finzi, J. Donato, M. Sucupira, G. De Nucci, J. Chromatogr. B 824 (2005) 30.
[19] P. Edder, S. Vargas, D. Ortelli, C. Corvi, Clin. Chem. Lab. Med. 41 (2003) 1608.
[20] A. Leitner, P. Zöllner, W. Lindner, J. Chromatogr. A 939 (2001) 49.
[21] Uhlig (2006): Personal communications.
[22] B. Jülicher, P. Gowik, S. Uhlig, Analyst 123 (1998) 173.
[23] Analytical Method Committee, Analyst 120 (1995) 2303.
[24] EURACHEM, EURACHEM/CITAC Guide, 2nd edn. (2000), ISBN 0 948926 15 5.