



Tandem and High Resolution Mass Spectrometry: a tool for food safety

Analytical performance criteria: current discussions in EU

Petra Gowik

EURL/NRL Berlin for Residues
crivetdrug@bvl.bund.de



- **Current development**
- **EU level (DG SANCO, EURLs, NRLs)**
 - **Multi-residue methods (definition – purpose)**
 - **Techniques**
 - **Criteria**
 - **Validation**
- **Codex level (CCRVDF)**
 - **Multi-residue methods (definition – purpose)**
 - **Techniques**
 - **Criteria**
 - **Validation**

Revision of Regulation (EC) No. 882/2004 and Directive 96/23/EC

- **Special focus on antibiotics is conceivable**
- **Changes in the requirements of substances/matrices to be tested for**
- **Discussion on MRPLs, RPAs**
- **Risk based approach**
- **How many samples/substances/matrices will be fixed**
 - ➔ **Frequency of sampling?**
 - ➔ **Frequency of analyses?**
 - ➔ **Strategy to be followed?**
 - ➔ **Kind of analytical technique to be applied?**

Validation and proficiency

- Transposition of the guidelines on validation of screening methods
- P(erformance) I(ndicator) specifications for PTs
- Multiannual evaluation of the PTs of the EURL Berlin



Multi-residue methods

- Definition of the term multi-residue methods
- Techniques
- Purpose
- Validation requirements for multi-residue methods especially for screening methods
- Performance criteria for multi-residue methods



COMMUNITY REFERENCE LABORATORIES RESIDUES (CRLs)
20/1/2010

**GUIDELINES FOR THE VALIDATION OF SCREENING METHODS
FOR RESIDUES OF VETERINARY MEDICINES
(INITIAL VALIDATION AND TRANSFER)**



Table of Contents

1.	Scope	1
2.	Definitions	1
2.1.	Regulatory/Action Limit for validation purposes	1
2.2.	Screening Target Concentration	1
2.3.	Detection capability $CC\beta$	2
2.4.	Cut-Off Level.....	2
2.5.	"Negative Control" (Blank matrix) Samples.....	2
2.6.	"Screen Positive Control" Samples.....	2
2.7.	Test Matrix	2
2.8.	Standard Operating Procedure (SOP) for the method	2
2.9.	Initial validation	3
2.10.	Abridged validation	3
2.11.	Screening Methods.....	3
3.	Screening Method classification.....	3
3.1.	Classification by detection principle	3
3.2.	Classification by their degree of quantification.....	3
4.	Principles to be followed for the validation of screening methods.....	4
4.1.	Key requirements.....	4
4.2.	Choice of analytes used for validation and selectivity of the method	4
4.3.	Preparation of "simulated tissue" for validation of microbial growth inhibition tests	6
5.	Validation procedure	7
5.1.	Determination of specificity/selectivity and detection capability $CC\beta$ according to classical approach	7
5.2.	Determination of specificity/selectivity and detection capability $CC\beta$ according to alternative matrix-comprehensive approach.....	10
6.	Transfer of screening methods between laboratories.....	10
6.1.	Conditions of transfer.....	10
6.2.	What is included in the abridged validation according to the classical concept.....	11
6.3.	What is included in the abridged validation according to the alternative concept.....	11
7.	Continuous verification.....	12
7.1.	Quality control (QC) samples.....	12
7.2.	Proficiency tests	13
8.	Validation report	13
9.	References.....	13
10.	Annexes	15

Guidelines

$CC\beta < \text{Rec Conc or MRL}$

Quantitative, semi-quantitative, qualitative

Orthogonal factorial design



Exemplary experimental design

Settings	Run	Systematic factors			Random factors			Concentration levels [CFU/g]			
		Thawing	Incub. time	Reading	Casting temp.	Nutrient media	Operator	500	50	0	0
1	1	1	-1	1	1	1	1				
2	1	-1	1	-1	1	1	1				
3	2	1	-1	-1	2	2	2				
4	2	-1	1	1	2	2	2				
5	3	1	-1	1	1	3	1				
6	3	-1	1	-1	1	3	1				
7	4	1	-1	-1	2	4	2				
8	4	-1	1	1	2	4	2				

Systematic factor: Thawing

- 1: 3 h at room temperature
-1: overnight at +2 to +4°C

Systematic factor: Incubation time

- 1: 24 h
-1: 18 h

Systematic factor: Reading

- 1: immediately after incubation (as described in standard)
-1: after storage in refrigerator over weekend

Random factor: Casting temperature

- 1: water bath at 47°C
2: water bath at 44°C

Random factor: Nutrient media

nutrient media from 4 different producers are to be used

Random factor: Operator

- at least 2 persons should take turns in processing the samples (up to 4 persons if possible)
1: person A
2: person B

FF.PT.2

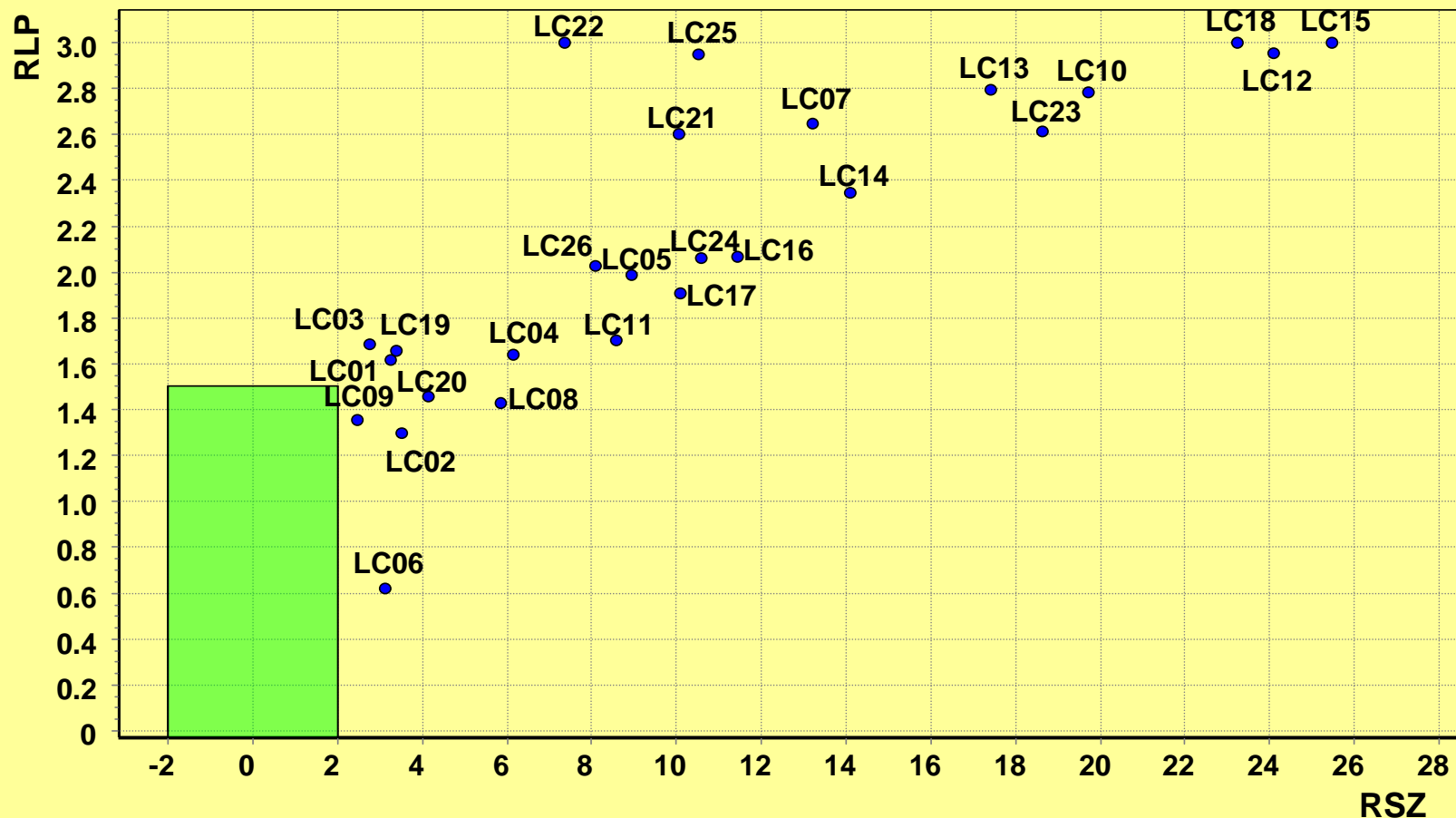
Grading addressing the **complexity of each PT** with corresponding justification

- 1 - easy matrix, one single easy analyte whose identity is shared with the participants, no incurred materials
- 2 - more challenging analytes and matrices, usually from an agreed list
- 3 - more complex analytes whose identity is not disclosed to the participants, innovative substances, a mix of analytes either in the same matrix or in different matrices, incurred materials

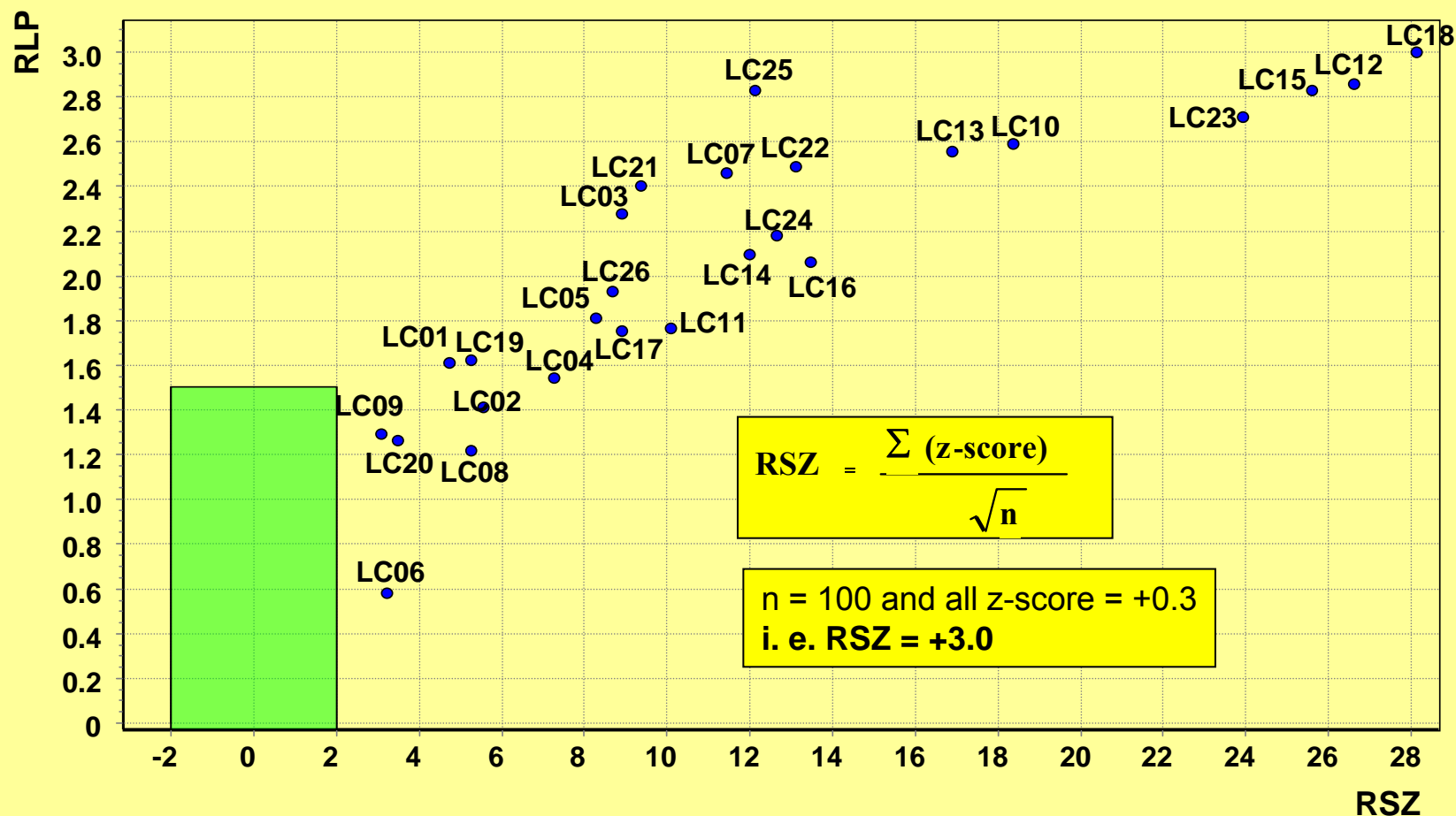
FF.PT.3

Average rates of NRL success
(share of NRLs that are expected to meet all the test thresholds) in relation to PT's grade of complexity and methods and activities to ensure **follow-up of poor results***

Interlaboratory studies 2002 up to 2008 (incl. all false negative) (NSAIDs, Coc.(2x), β -Ago., Nitroimi. and Anth. (2x) – 72 analytes)

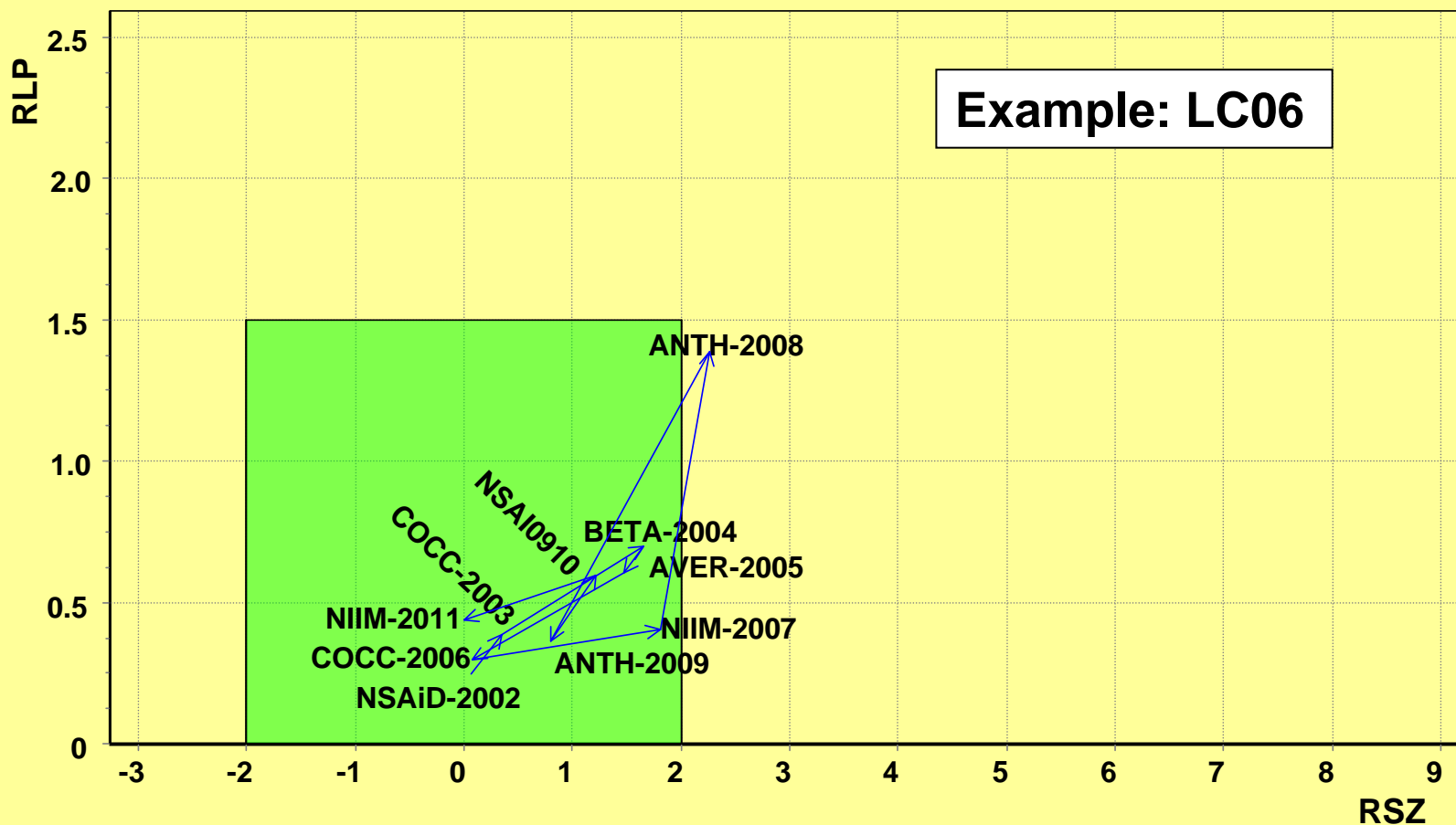


Interlaboratory studies 2002 up to 2011 (incl. all false negative) (NSAIDs(2x), Coc.(2x), β-Ago., Nitroimi.(2x) + Anth. (3x) – 100 analytes)



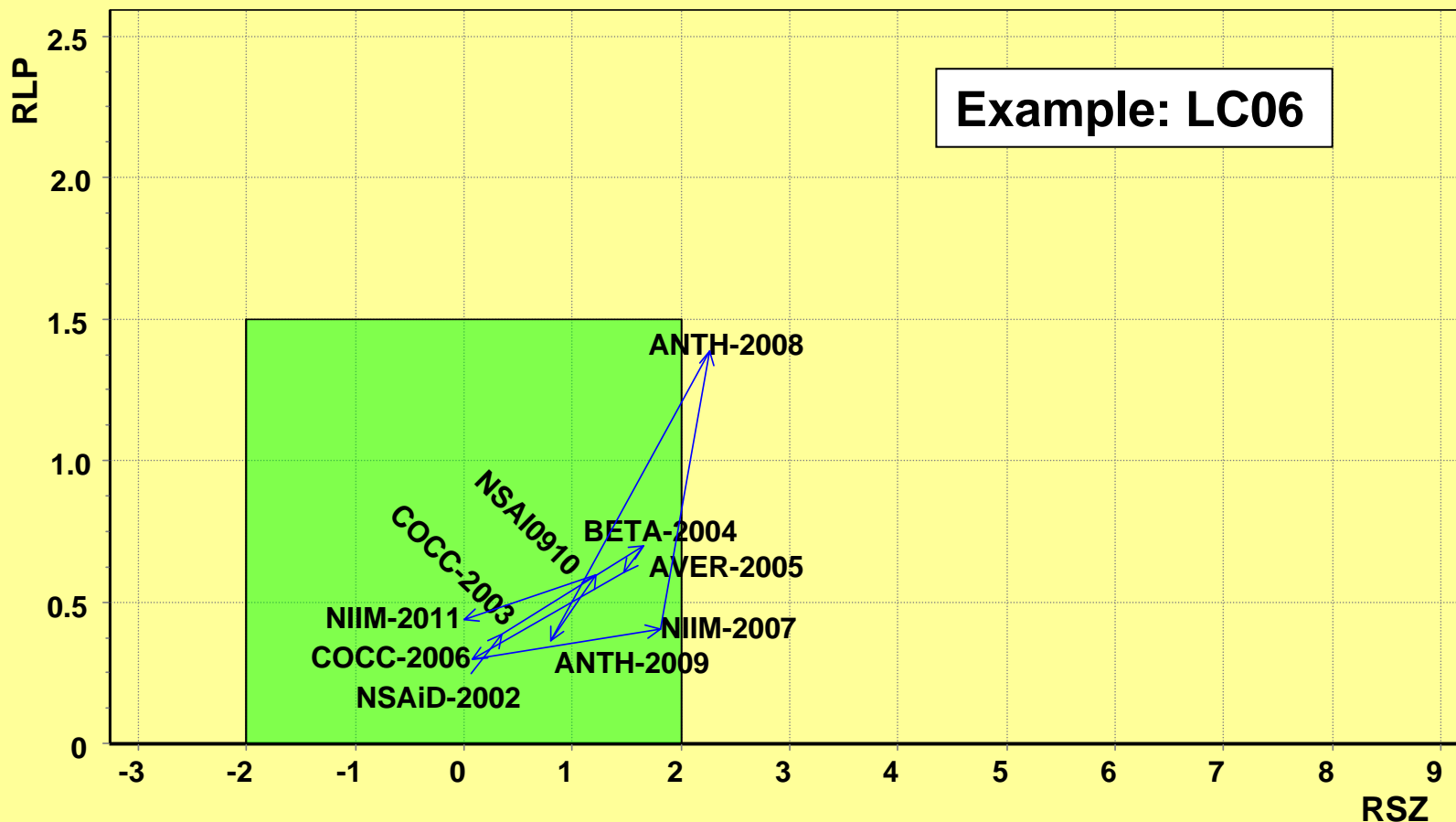
Combination scores for one laboratory (2002 – 2011)

(without false negative results)

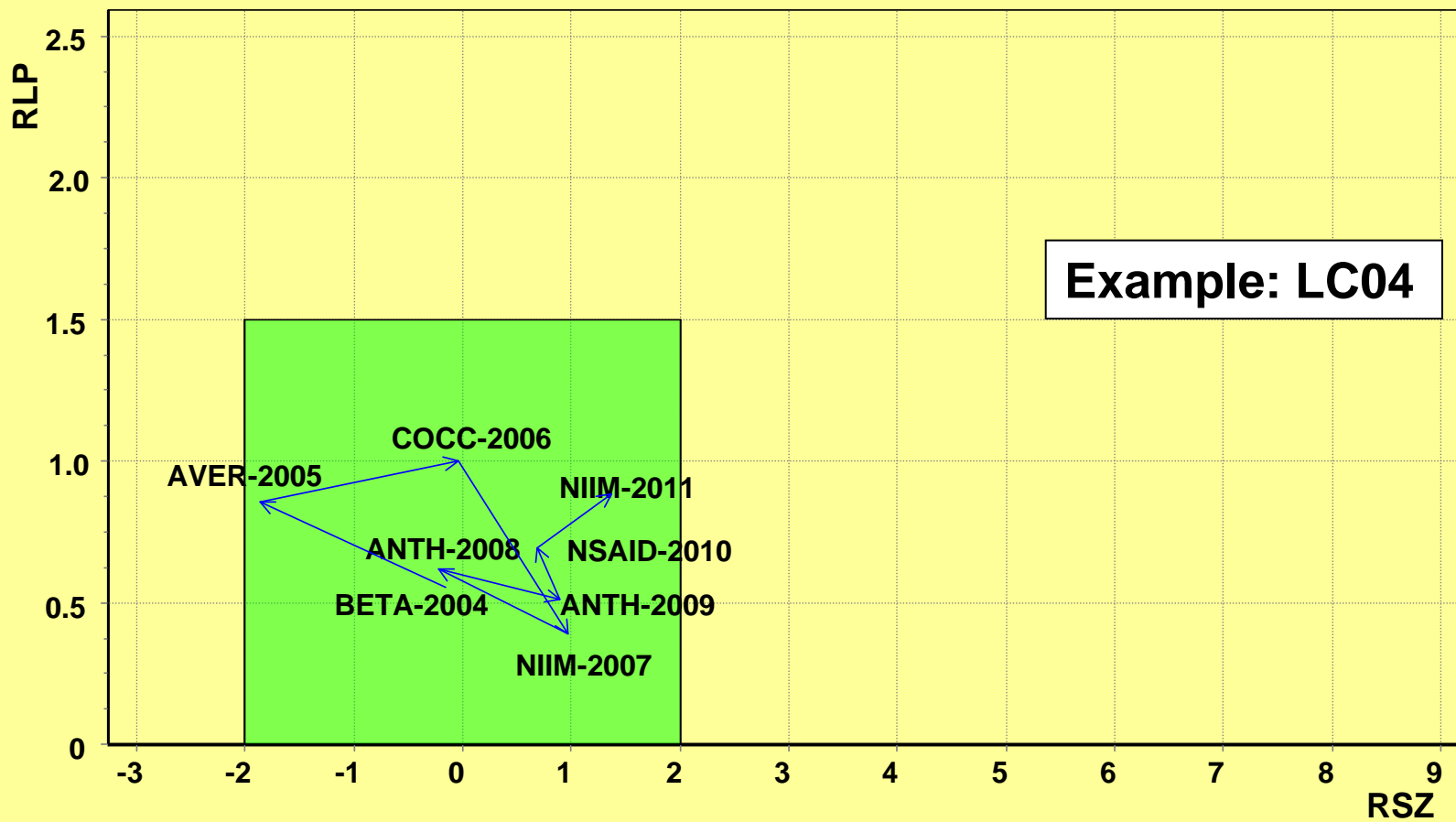


Combination scores for one laboratory (2002 – 2011)

(incl. negative results)

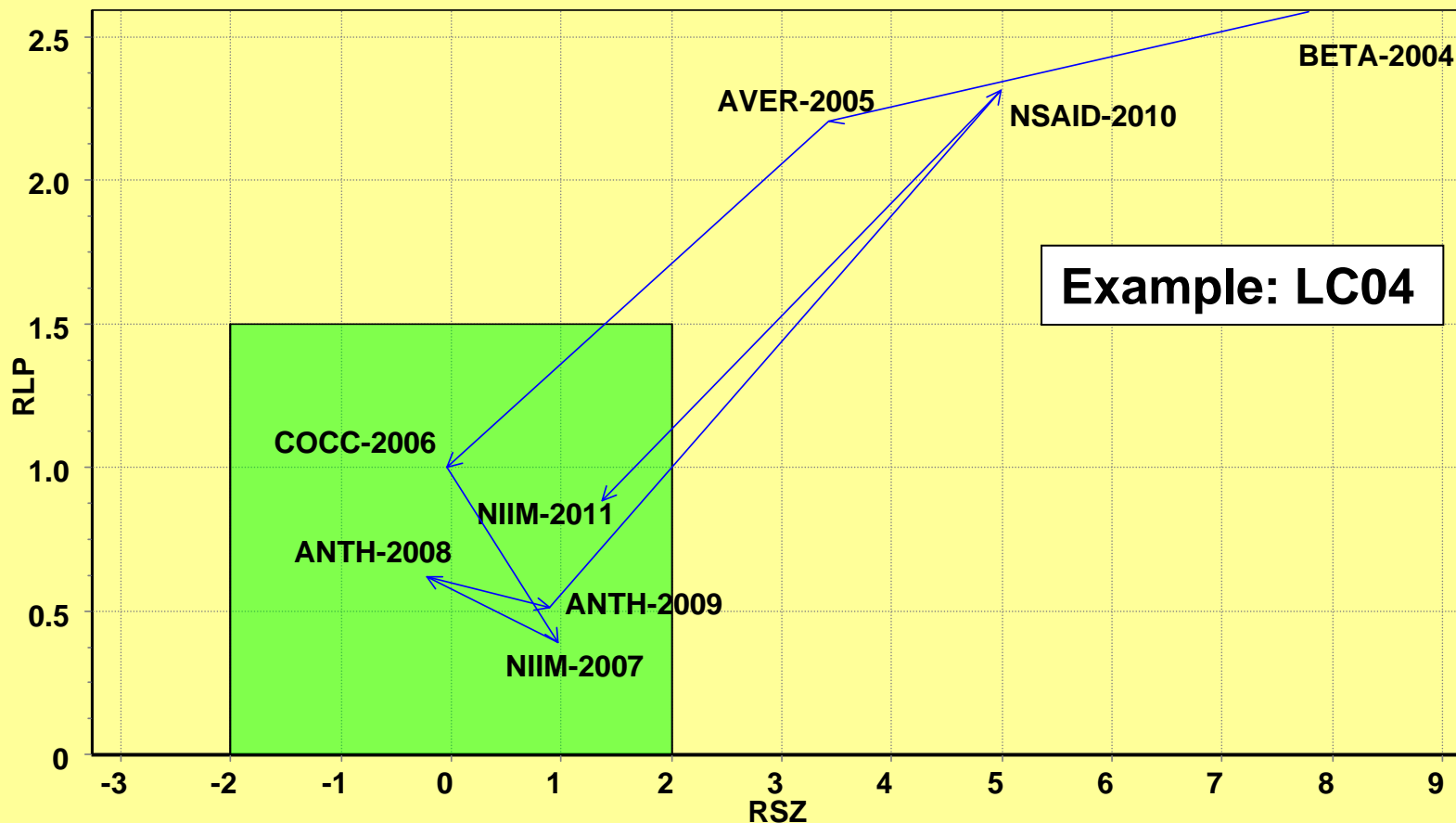


Combination scores for one laboratory (2004 – 2011) (without false negative results)



Combination scores for one laboratory (2004 – 2011)

(incl. all false negative results)



PT providers use different evaluation and assessment schemes

Assessment is not harmonised

Results of the different PT providers are not comparable among each other

For the sake of comparability assessment criteria should be harmonised

Discussion on the EURL meeting in June 2012

On German national level the accreditation body published criteria to be followed for the pesticides' area just recently



Anforderungen an Laboratorien im Gesundheitlichen Verbraucherschutz – Wirkstoff–Multimethoden zur Pestizidanalytik in Lebens- und Futtermitteln

71 SD 4 012 | Revision: 1.2 | 17.09.2012

Geltungsbereich:

Diese Regel legt verbindliche Anforderungen und Rahmenbedingungen fest, unter denen Prüflaboratorien für den Bereich der Pestizidanalytik in Lebens- und Futtermitteln akkreditiert werden.

Performance Indicators/PT

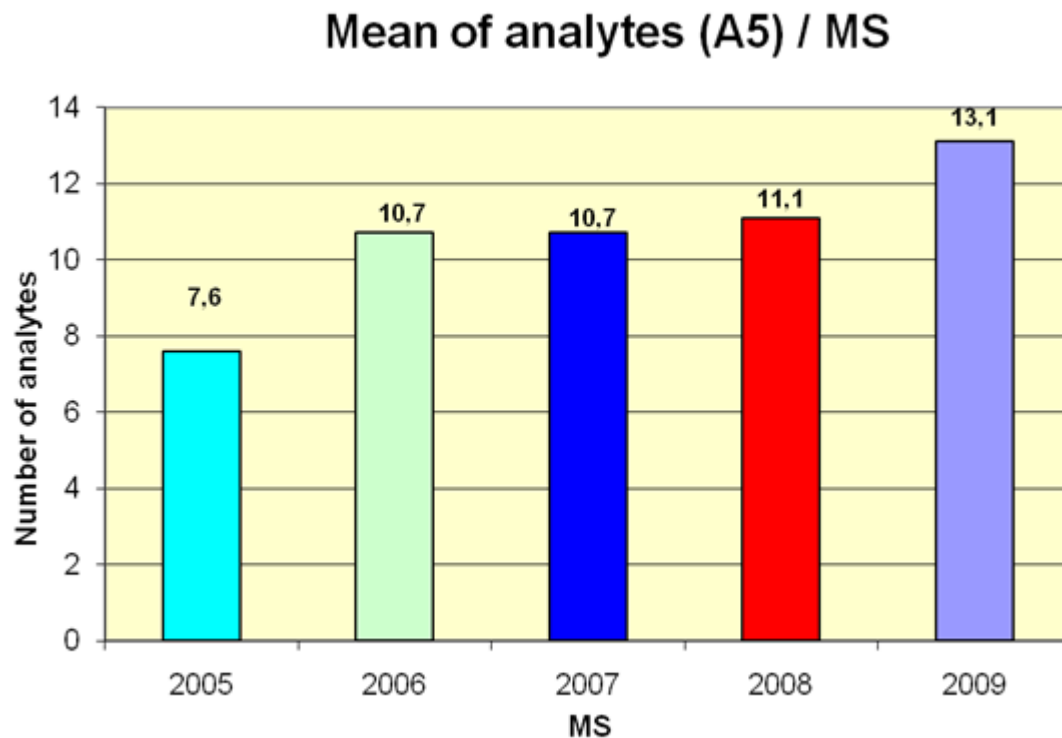
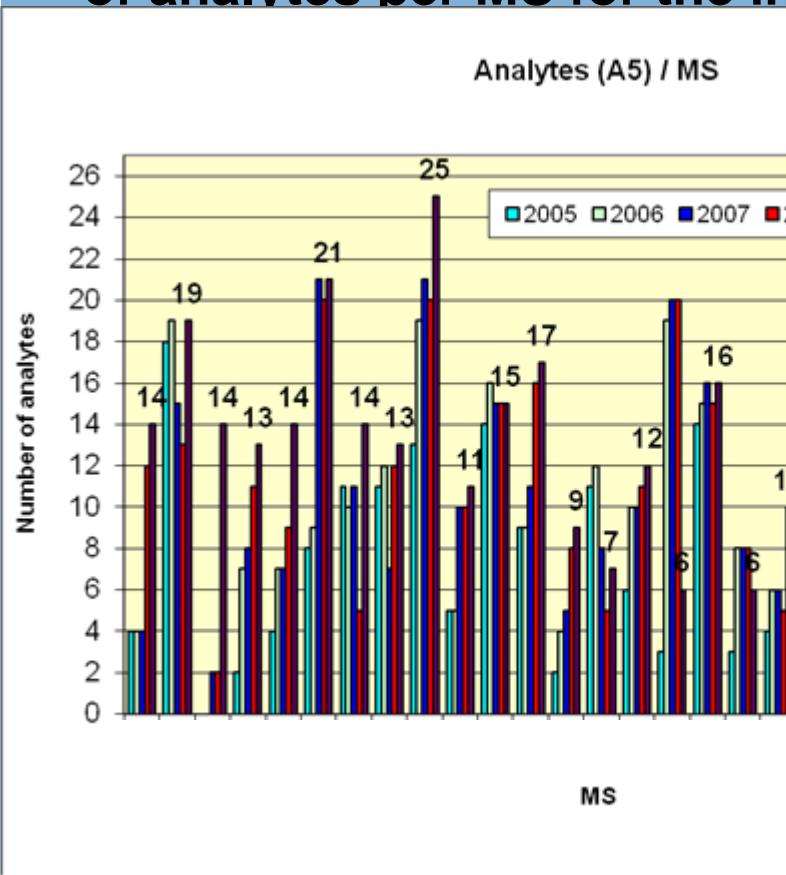
„... Eine Eignungsprüfung gilt nur dann als erfolgreich bestanden, wenn keine falsch-positiven Wirkstoffe angegeben und mindestens 75% der quantitativ angegebenen Wirkstoffe mit einem z-Score bis zu ± 2 bestimmt wurden. Falsch-negative Befunde gehen mit hohem negativen z-Score in das Bewertungsschema ein. ...“

http://www.dakks.de/doc_pl-umwelt

Questions:

- **Definition of the term multi-residue methods**
- **Techniques**
- **Purpose**
- **Validation requirements for multi-residue methods especially for screening methods**
- **Performance criteria for multi-residue methods**

- **Slow but continuous evolution of establishment of multi-residue methods**
- **Reports to DG SANCO on NRCP evaluation included the enumeration of analytes per MS for the individual substance groups**



- For all substance groups the EURL developed and will develop multi-residue methods, also for provision to other official labs
- EURL always advocated the use of MRM
- Meanwhile also a multi-substance group method is implemented at EURL (presented at this WS) including B1, B2a,b,e and will be extended
- Since years there is a trend to establish multi-substance groups methods of up to 300 analytes (screening)
- This trend produces new questions and requirements to be answered and established, respectively
- High workload for validation

Questions *(list not closed)*

- What is a possible and reasonable application?

Keywords: screening - confirmation

- What kind of requirements /criteria are necessary to guarantee a sufficient performance?

Keywords: identification, precision, reliability, quantification

- How should validation look like?

Keywords: efficiency, representative analytes, transfer of methods, standardised methods?

- Do we need a new strategy?

Keywords: fully validated confirmatory methods in stock? Or validation on demand?

10th anniversary of 657!

- **Criteria layed down in 657/2002 are applicable to multi-residue methods**
- **Definition of multi-residue, how many analytes?**
- **For >100 analytes it is difficult to realise and fulfill all requirements concerning criteria and validation in reasonable time frames**
- **Might have an Effect on requirements layed down in 96/23 concerning sampling and number of samples per substance group, to be clarified if substance group comprehensive analyses are in line with these requirements**
- **Indispensable for control of AB**
- **Do we need new criteria for new techniques?**



Questionnaire to NRLs

The aim and the background of this survey is the identification of possible changes in the performance criteria of the validation requirements for multi-residue methods with very high numbers of analytes.

In the following you will be asked a few questions regarding multi-residue methods. The term might be understood in different ways. Therefore a definition is presented for the particular purpose of this questionnaire.

Multi-residue methods in this questionnaire are understood to mean methods including around 50 and more analytes comprising different substance groups; they may be able to cope with several matrices. Most likely generic methods are meant and used for this purpose. Emphasis is certainly put on screening methods.

Questionnaire to NRLs

Do you use multi-residue methods for

Which methods are used in your lab for multi-residue purposes?

How many substance groups are included?

How many substances are included?

How many matrices are covered?

If LC-QTOF or LC-Orbitrap Instruments are used: How do you identify the substances? (product ion spectra – exact mass?)

As NRL do you have an overview whether routine labs apply Multi-residue Methods?

If yes, do you have an overview how routine labs validate?



Questionnaire to NRLs

How do you validate multi-residue methods for a high number of analytes (Screening)?

Number of samples

number of concentration levels

number of parallel samples

acc. to CRL guidelines 2010 classical approach or alternative approach

How do you validate multi-residue methods for a high number of analytes (Confirmation)?

Number of samples

number of concentration levels

number of parallel samples

acc. to CD 2002/657/EC classical approach or alternative approach

Which Validation parameter do you determine (Screening)?

Which Validation parameter do you determine (Confirmation)?



Questionnaire to NRLs

Do you use multi-residue methods for

Which methods are used in your lab for multi-residue purposes?

How many substance groups are included?

How many substances are included?

How many matrices are covered?

If LC-QTOF or LC-Orbitrap Instruments are used: How do you identify the substances? (product ion spectra – exact mass?)

As NRL do you have an overview whether routine labs apply Multi-residue Methods?

If yes, do you have an overview how routine labs validate?



Function	Mass resolution (FWHM) [G1]	Mass accuracy (mDa)	Remarks
Screening	$\geq 10,000$	± 50	Relative retention time $\leq 2.5\%$
Confirmation	$\geq 10,000$	≤ 5	1.5 [G2] IPs per ion or product-ion, min. 1 ion ratio, Relative retention time $\leq 2.5\%$
HR confirmation	$\geq 20,000$	≤ 5	2 IPs per ion or product-ion, min. 1 ion ratio, Relative retention time $\leq 2.5\%$
MS/MS identification of unknowns	$\geq 10,000$	≤ 5	Confirm postulated structure by NMR [G3] and/or confirm accurate masses at mass resolution $\geq 70,000$ [G4] (FWHM)

Proposal for additional LC-MS identification criteria to be supplemented to the 2002/657/EC (6); adapted from Nielsen et al (20).

[\[G1\]](#) at which mass? this parameter is mass-sepdependent

[\[G2\]](#) what was the reason for appointing 1.5 points to the parent ion now? 1.0 point is sufficient.

[\[G3\]](#) the postulation to confirm the analyte by NMR is too sophisticated

[\[G4\]](#) 70,000 (FWHM) is binding to orbitrap or FT. Do think that should be really required? My colleague said that to his and others experiences an average of 50,000 is sufficient .



CCRVDF

DRAFT CCRVDF GUIDELINES FOR THE VALIDATION OF MULTI-RESIDUE ANALYTICAL METHODS FOR THE SCREENING AND DETERMINATION OF VETERINARY DRUG RESIDUES IN MATRICES FROM FOOD PRODUCING ANIMALS



May 2011:	Discussion with NRLs at EURL workshop
10/05/11:	Comments sent to eWG
Oct. 2011:	Further comments to the revised draft as of Oct. 2011
13/02/12:	Further comments to the revised draft as of Dec. 2011
15/03/12:	Draft EU comments sent to MS for comments
17/04/12:	Revised Draft EU comments distributed to MS
25/04/12:	Meeting of MS at Council Secretariat for final discussion Forwarded to CCRVDF secretariat
7-11/05/12:	20th session of CCRVDF at Puerto Rico



**Our view on questions raised in the presentation by
Jack Kay**

**Performance criteria
for multi-residue analytical methods (MRM)**

Question 1

Can the 5% (False Positive/False Negative rates) for single analyte methods, safely be extended to MRMs?

Remarks:

The 5% criteria can be extended to MRMs for MRL-substances.

Why no criteria for non-authorized substances (1%?)

Analytes which don't fulfil these criteria should be able to be included in the methods for screening and the false positive and the false negative rate have to be controlled and confirmed regularly.



Question 2 (Table 1)

Is it recommended that methods used to support Codex MRLs should meet the performance standards for trueness and precision listed in Table 1?



Topics discussed at EURL Workshops

Concentration	Coefficient of variability (CV)				Trueness
	Repeatability (within-laboratory, CV _A)	Repeatability (within-laboratory, CV _L)	Reproducibility (between-laboratory, CV _A)	Reproducibility (between-laboratory, CV _L)	Range of mean % recovery*
(µg/kg)	(%)	(%)	(%)	(%)	(%)
≤ 1	35	36	53	54	50–120
1 to 10	30	32	45	46	60–120
10 to 100	20	22	32	34	70–110
100 to 1 000	15	18	23	25	70–110
≥ 1 000	10	14	16	19	70–110

Codex Draft

Mass fraction	Coefficient of variability (CV)				Minimum trueness
			Reproducibility (CV)	Mass fraction	Range %
(µg/kg)			(%)	(µg/kg)	(%)
1			(*)	≤ 1	-50 to +20
10			(*)	> 1 to 10	-30 to +10
100			23	≥ 10	-20 to +10
≥ 1 000			16		
* <Horwitz-SD and ALARA					

CD 657/2002

Question 2 (Table 1)

Remarks:

Why two values for repeatability and reproducibility (CV_A and CV_L)?

→ How to evaluate this values (by separate analyses? → very work intensive)

Are the fixed values for reproducibility (Horwitz-SD) for concentrations lower than 100 µg/kg still up to date? → Thompson, Analyst 2002

Which value for instance for 100 µg/kg (23% or 32%)?

No fixed values for complete conc. ranges – better is: “at least Horwitz-SD or ALARA”

What kind of recovery is meant? → equivalence (internal standards or matrix calibration)

absolute recoveries are not of importance as long as sensitivity and precision are sufficient

[Recovery rates are slightly different from 657/2002/EC]

Question 3

Are the values in this table still acceptable and appropriate?

Relative ion intensity (% of base peak)	GC-MS (EI) (relative)	GC-MS (CI), GC- MS/MS, LC-MS, LC-MS/MS (relative)
(%)	(%)	(%)
> 50	≤ 10	≤ 20
20–50	≤ 15	≤ 25
10–20	≤ 20	≤ 30
≤ 10	≤ 50	≤ 50

Draft

=

657/2002

Question 3 (Table 2 – draft)

Are the values in this table still acceptable and appropriate?

***Consensus in all discussion rounds.
up to now no changes***

Remarks to the text of draft (1)

- *Point 16: → in-house or within-lab reproducibility is not mentioned*
- *Point 20: → within-laboratory variation = repeatability is not correct*
- *Point 23: → the weakest relevant response of the analyte plus three times its standard deviation → is not up to date for some instrumental methods*
- *Point 25: → terms Decision Limit ($CC\alpha$) and Detection Capability ($CC\beta$) are not an alternative to using LOD and LOQ*
 - $CC\alpha$ and $CC\beta \neq$ LOD and LOQ
 - the definition of $CC\beta$ in the glossary is not OK

Remarks to the text of draft (2)

- *Point 34:* *Internal standards are not mentioned in the draft.*
- *Point 41:* *“Analyte stability during analysis must be established for both standards and analyte in the presence of sample material, during processing through the complete analysis for all methods used in a regulatory control programme and for typical conditions of storage while a sample is awaiting analysis.” → should be simplified*
- *Point 46:* *residue values below 100 µg/kg can not be expressed with one significant figure only
→ e. g. what does it mean in case of Meloxicam (MRL: 15 µg/kg) or in case of Diclofenac (MRL: 0.1 µg/kg)?*