Anal Bioanal Chem (2013) 405:7875–7884
DOI 10.1007/s00216-013-7070-3

RESEARCH PAPER

An inter-laboratory validation of a multiplex dipstick assay for four classes of antibiotics in honey

Katharina Heinrich · Roy Macarthur · Christoph von Holst · Matthew Sharman

Received: 28 January 2013 / Revised: 8 May 2013 / Accepted: 16 May 2013 / Published online: 3 July 2013

Abstract In this paper, we report the inter-laboratory validation (ILV) of a recently developed indirect competitive multiplex dipstick (Bee4sensor®) which is capable of the simultaneous detection of residues of some of the most frequently detected antibiotic residues in honey: sulfonamides, tylosin, fluoroquinolones and chloramphenicol. The multi-sensor dipstick can be interpreted via visual observation or by an instrumental measurement of four test lines. Statistical analysis of the ILV data demonstrated that the multi-sensor can reliably detect the presence of sulfathiazole at 25 μg kg^{-1} and tylosin at 10 μg kg^{-1}, which fully meet the 'recommended concentrations' of the EU. Ciprofloxacin and chloramphenicol can be detected at 25 and 5 μg kg^{-1} in honey, respectively. Whilst the concentration for chloramphenicol is above the EU minimum required performance limit of 0.3 μg kg^{-1}, this part of the multiplex test may still be of use to both the industry and enforcement authorities, to provide an early warning of contaminated honey. The estimated false-negative and false-positive rates for this easy-to-use and robust assay were less than 5 %.

Keywords Honey · Antibiotics · Screening method · Bioassay · Drug monitoring · Food/beverages

Published in the topical collection *Rapid Detection in Food and Feed* with guest editors Rudolf Krska and Michel Nielen.

K. Heinrich (✉) · R. Macarthur · M. Sharman
The Food and Environment Research Agency (Fera), Sand Hutton, York, YO41 1LZ, UK
e-mail: katharina.heinrich@fera.gsi.gov.uk

C. von Holst
European Commission, Joint Research Centre, Institute for Reference Materials and Measurements, Retieseweg 111, 2440 Geel, Belgium

Introduction

Pollination by bees is essential for the sustainability of many sectors of agricultural production; therefore, maintaining the health of bee colonies is of great economic importance. Bee colonies are susceptible to a number of infestations and diseases including varroa and foulbrood, but worldwide there are relatively few drugs legally permitted for control [1]. Although some countries, e.g. the USA, have authorised the use of the antibiotic tylosin to treat American foulbrood [2], the EU has yet to authorise antimicrobial-based veterinary medicines for the treatment of bees. Between 2007 and 2012, there were, however, approximately 100 alerts for honey and royal jelly reported on the European Commission's Rapid Alert System on Food and Feed [3]. Of these, some of the most frequent types of antimicrobials detected were sulfonamides (e.g. sulfadiazine, sulfathiazole, sulfamethoxazole), aminoglycosides (streptomycin), fluoroquinolones (e.g. ciprofloxacin) and chloramphenicol.

Whilst there are no EU maximum residue limits (MRLs) for any antimicrobials in honey, a minimum required performance limit (MRPL) for chloramphenicol has been set at 0.3 μg kg^{-1} [4]. In addition, the EU Reference Laboratories have published a number of 'recommended concentrations' (RCs) in honey ranging from 20 μg kg^{-1} for tylosin to 50 μg kg^{-1} for sulfonamides [5]. These RCs are related to the screening target concentration (STC). The STC is the concentration at which a screening test categorises the sample as 'screen positive' (potentially non-compliant) and prompts a confirmatory test by, for example, liquid chromatography–tandem mass spectrometry (LC-MS/MS). For analytes for which MRLs have not been established, the 'STC should wherever possible be at or less than the

recommended concentrations as described in the CRL Guidance Paper' [6]. For other non-authorised analytes (veterinary medicines) in honey, e.g. fluoroquinolones, the STC is at the discretion of the laboratory conducting the work. For this study, the STCs used were as follows: 50, 20, 50 and 10 μg kg^{-1} for sulfathiazole, tylosin A, ciprofloxacin and chloramphenicol, respectively.

Concern about the occurrence of residues of veterinary medicines in honey has resulted in the need for rapid methods to monitor for antibiotic residues from a variety of classes including sulfonamides, macrolides, fluoroquinolones and tetracyclines. Whilst a sensitive and rapid method based on dipstick technology already exists for tetracyclines in honey (TETRASENSOR® kit, Unisensor, Belgium) [7], and there are a number of commercial ELISA products for streptomycin and tylosin [8], there is a dearth of rapid methods which offer a wide coverage of the antimicrobial classes of interest.

In 2007, the European Commission Framework Programme (FP7) commissioned a new project called Contaminants in Food and Feed: Inexpensive Detection for Control of Exposure (Con*ff*idence, www.conffidence.eu). The aim of the project was to develop and validate faster and more cost-efficient methods for the detection of a wide range of chemical contaminants/residues in both food and feed. One of the outcomes from this project is a multiplex dipstick assay (Bee4sensor® [9, 10]) that could be employed for the simultaneous screening of sulfonamides, fluoroquinolones, tylosin and chloramphenicol, possibly present in honey. Data from this method could be determined in two ways:

(a) Visually, in which the colour intensities of the individual test lines on the dipstick (for sulfonamides, fluoroquinolones, tylosin A and chloramphenicol) are compared to the intensity of a control line
(b) Instrumentally, in which a Readsensor® is used to generate a numerical value (ratio) for the colour intensity

Initially, this new method was subjected to a within-laboratory validation according to EU requirements (Commission Decision 2002/657/EC [11]) for a wide range of analytes from both the sulfonamide and fluoroquinolone classes of antimicrobials, as well as tylosin and chloramphenicol. Subsequently, to ensure robustness of this new method, a small-scale inter-laboratory validation (ILV) exercise was undertaken using 'representative analytes' [6] of sulfathiazole, ciprofloxacin, tylosin and chloramphenicol. This paper reports the results of this external validation study by seven control/industry laboratories. This validation trial was conducted in two stages. The laboratories familiarised themselves with the method by analysing known positive and negative control samples, before testing a number of randomly coded blind samples, containing spiked and incurred residues of the target antimicrobial compounds.

Materials and methods

Materials and reagents

The analytical standards of sulfathiazole, tylosin A, ciprofloxacin (hydrochloride) and chloramphenicol were purchased from QMX Laboratories (Thaxted, UK). Chemicals and solvents were of analytical grade. Deionised double-distilled (or equivalent) water was used throughout.

The multi-sensor (Bee4sensor®) kits for honey, microwells and dipsticks (lot 110505), were provided by Unisensor s.a. (Liège, Belgium) and used during the inter-laboratory validation. Bulk buffer solutions were prepared separately to cater for all the participants of the trial.

Stock and working standard preparation

Standard stock solutions of 1,000 μg ml^{-1} were made in methanol and kept at −20 °C for up to 6 months. Intermediate solutions of 100 or 10 μg ml^{-1} were prepared on a monthly basis and also kept at −20 °C, while mixed spiking solutions were prepared on a weekly basis and stored at 4 °C.

Instrumentation

The Heatsensor® equipment and the Readsensor® (for instrumental reading) were supplied to each laboratory, by Unisensor s.a. (Liège, Belgium). This Readsensor® version (Firmware 2.2.13) can be used in stand-alone mode or connected to a computer. In this study, the device was used in computer mode and was operated with specially designed software (either LF control, version 3.01.01; or LF Studio, version 3.3.4). To ensure correct performance of the instrumental system, the Readsensor® equipment was calibrated by each participant in the ILV immediately before each part of the study with a 'reference check & control dipstick' as recommended by the provider.

Samples

Blind test material no. 1 ('blank')

A 'blank' honey was produced by stirring portions (previously tested negative using an ISO 17025-accredited LC-MS/MS procedure) of acacia, woodland, orange blossom, Australian honey (5:5:5:4 ratio) for 24 h. After mixing, aliquots (15 g) were potted and stored at ambient temperature.

Blind test material no. 2 ('0.5× STC')

The following protocol was employed to produce a honey containing sulfathiazole, tylosin A, ciprofloxacin and

chloramphenicol at 25, 10, 25 and 5 μg kg^{-1}, respectively. To 1,000 g of 'blank' honey were gradually added 0.25 ml of 100 μg ml^{-1} sulfathiazole and ciprofloxacin solutions, 0.1 ml of 100 μg ml^{-1} tylosin A and 0.5 ml of 10 μg ml^{-1} chloramphenicol, whilst the honey was being stirred continuously and then for a further 24 h at room temperature. Aliquots (15 g) were potted and stored at −20 °C.

Blind test material no. 3 ('1× STC')

To produce a honey containing 50/20/50/10 μg kg^{-1} of sulfathiazole/tylosin A/ciprofloxacin/chloramphenicol, the following protocol was employed. To 1,300 g of 'blank' honey were gradually added 0.65 ml of 100 μg ml^{-1} sulfathiazole and ciprofloxacin solutions, 0.26 ml of 100 μg ml^{-1} tylosin A and 0.13 ml of 100 μg ml^{-1} chloramphenicol, whilst the honey was being stirred continuously and then for a further 24 h at room temperature. Afterwards, aliquots (15 g) were prepared and stored using the same procedure as for test material no. 2

Other test materials (incurred)

Test material no. 4 (three honey samples containing known tylosin A concentrations) and test material no. 5 (incurred honeys with varied concentration of sulfonamides, ciprofloxacin and chloramphenicol blended with 'blank') were also provided, but for quality control purposes only, and results are neither presented nor discussed.

Homogeneity testing of test materials

Ten aliquots of each test material (2, 3, 4 and 5) were extracted in duplicate and analysed by LC-MS/MS, following ISO/IEC17025:2005-accredited procedures, validated to Commission Decision 2002/657/EC. These data showed that the test materials were homogenous (data not presented), according to ISO17043. The concentrations of the various analytes in incurred material 4 and 5 varied between 3.8 and 93.6 μg kg^{-1}.

Design of inter-laboratory study

This study was conducted in two stages at the beginning of 2012. Each laboratory was supplied with two protocols, one for the procedure of Bee4sensor® multiplex method and one for the protocol of the trial. During the pre-trial, the seven participating labs from five European countries had the opportunity to familiarise themselves with the method, analysing two honey samples of known residue concentration ('blank' and '1×STC').

For the main trial, six replicate samples of 'blank', '0.5× STC' and '1×STC' honey were analysed blind (i.e. all sample pots supplied to the participating laboratories were allocated a random code) over 3 days, together with two quality control samples (test material 4 and 5) on each day. This mini-trial was therefore designed to provide 42 data points per concentration level for the subsequent statistical evaluation.

Test protocol and interpretation of results

For the analysis of a honey sample, two aliquots (A and B) of 2.5 g each are required. The procedure used in this trial was as described in the kit insert. Briefly, one aliquot (A) is dissolved using acid hydrolysis, whereas the other aliquot (B) is dissolved in water. After liquid/liquid partitioning both aliquots with ethyl-acetate, the organic layers are evaporated to dryness under nitrogen. After reconstitution in a buffer, aliquots A and B are combined and applied to the well of a Bee4sensor® test kit for 5 min at 40 °C, as one sample extract. A dipstick is then incubated in this prepared well for 10 min at 40 °C (using the Heatsensor® device).

The Bee4sensor® kit is an indirect competitive-based assay in a dipstick format. The dipstick has five green capture lines, one for the control and one each for chloramphenicol (CAP), fluoroquinolones (QUINO), tylosin A (TYL) and sulfonamides (SULFA), which turn red (see Fig. 1) when in contact with the rising liquid.

For each test sample, the dipstick was assessed both visually and instrumentally via the Readsensor®. In both approaches, the colour intensity of the four capture lines of the analytes (test lines) is compared to the control capture line.

In the visual assessment, samples that 'screen positive' (potentially non-compliant) will have an absent or a reduced/weaker red colour (compared to the control line) at the test line. In the case of compliant (or 'negative') samples, the test line will show an increased red colour

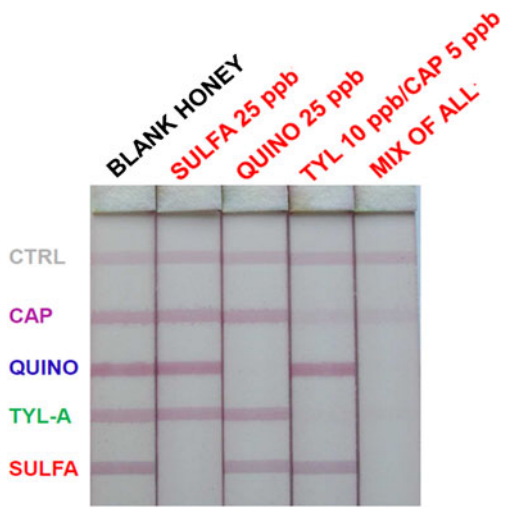

Fig. 1 Example of multiplex dipstick test lines for a blank honey sample and honey spiked at 0.5×STC

compared to the control line. If the control line fails to become visible, the result of the dipstick is recorded as invalid.

During the instrumental assessment, the intensity of the developed red colour at each of the capture lines is measured, and the results are expressed as the ratios of these measurements (test line versus control line). The software method was programmed with a certain set of ratios to enable objective interpretation of the colour intensity: Honey samples were reported (a) as free of antibiotics (i.e.

Table 1 Readsensor® results for unfortified (blank) honey

STC [%]	Lab	Day	ID	Sulfathiazole	Tylosin	Ciprofloxacin	Chloramphenicol
0	1	1	37	2.58	1.90	1.71	2.60
0	1	1	85	2.12	1.29	1.43	2.06
0	1	2	120	3.47	3.77	2.23	3.14
0	1	2	156	3.12	3.43	2.09	2.87
0	1	3	186	3.29	3.34	1.93	2.91
0	1	3	206	2.69	2.55	1.56	2.30
0	2	1	40	2.13	2.07	1.87	2.24
0	2	1	93	3.63	2.49	2.52	3.35
0	2	2	121	1.98	2.25	1.96	2.23
0	2	2	159	2.09	2.35	1.86	2.29
0	2	3	189	2.58	1.79	1.45	2.13
0	2	3	211	2.24	2.37	2.03	2.21
0	3	1	48	2.13	2.23	1.86	2.25
0	3	1	95	2.08	2.29	1.78	2.17
0	3	2	126	1.96	2.09	1.58	1.94
0	3	2	160	2.00	2.06	1.56	1.84
0	3	3	192	2.07	2.11	1.59	2.14
0	3	3	214	2.23	2.42	1.83	2.29
0	4	1	55	2.96	2.56	1.53	2.59
0	4	1	104	2.68	2.20	1.59	2.54
0	4	2	131	3.25	3.17	1.74	2.70
0	4	2	161	2.92	3.20	1.74	2.66
0	4	3	197	3.19	2.43	1.81	2.57
0	4	3	217	2.92	2.31	1.59	2.40
0	5	1	58	3.91	3.62	1.92	2.84
0	5	1	105	4.02	3.73	1.88	2.84
0	5	2	134	4.00	4.32	2.52	3.32
0	5	2	168	3.63	3.63	2.27	3.06
0	5	3	198	3.12	3.38	2.01	2.97
0	5	3	219	3.21	3.53	2.19	3.03
0	6	1	59	2.24	2.52	2.18	2.36
0	6	1	107	2.38	2.47	2.03	2.71
0	6	2	139	2.05	2.72	2.75	2.87
0	6	2	171	2.48	2.50	2.48	2.61
0	6	3	199	2.40	2.56	2.40	2.84
0	6	3	221	2.49	2.16	2.13	3.43
0	7	1	65	3.93	3.51	2.58	3.31
0	7	1	113	3.56	3.44	2.58	3.54
0	7	2	140	4.61	4.31	2.76	3.62
0	7	2	174	3.26	3.22	2.19	3.10
0	7	3	200	3.55	3.51	2.24	3.19
0	7	3	223	3.84	3.41	2.04	3.04

'compliant') with a ratio >1.1, (b) to contain trace antibiotic residues ('low positive') with a ratio of 0.9 to 1.1, and (c) to contain a higher concentration(s) of antibiotic ('positive') with a ratio <0.9. The visual assessment of the dipsticks was performed after the instrumental reading, but without looking at the instrumental results to avoid any bias.

Results and discussion

Visual assessment

Data from the qualitative results (visual assessment) were used to estimate an upper confidence interval for average

Table 2 Readsensor® results for honey fortified at 50 % of the screening target concentration

STC [%]	Lab	Day	ID	Sulfathiazole	Tylosin	Ciprofloxacin	Chloramphenicol
50	1	1	5	0.00	0.00	0.00	0.56
50	1	1	26	0.00	0.00	0.00	0.00
50	1	2	60	0.53	0.97	0.29	0.71
50	1	2	91	0.47	0.78	0.25	0.90
50	1	3	136	0.38	0.48	0.00	0.77
50	1	3	170	0.77	0.69	0.00	0.46
50	2	1	8	0.35	0.53	0.38	0.82
50	2	1	29	0.34	0.58	0.28	0.87
50	2	2	61	0.06	0.30	0.07	0.86
50	2	2	94	0.07	0.39	0.26	0.95
50	2	3	138	0.00	0.30	0.17	0.79
50	2	3	175	0.00	0.29	0.29	0.75
50	3	1	9	0.00	0.37	0.33	0.77
50	3	1	34	0.30	0.58	0.37	0.88
50	3	2	68	0.12	0.31	0.31	0.83
50	3	2	106	0.25	0.45	0.22	0.80
50	3	3	141	0.31	0.40	0.35	0.99
50	3	3	194	0.23	0.24	0.25	0.73
50	4	1	12	0.74	0.00	0.00	0.61
50	4	1	35	0.73	0.00	0.00	0.74
50	4	2	71	0.56	0.34	0.24	0.51
50	4	2	111	0.81	0.51	0.33	0.58
50	4	3	150	0.38	0.38	0.29	0.69
50	4	3	196	0.42	0.22	0.25	0.82
50	5	1	14	0.77	1.31	0.00	0.79
50	5	1	38	0.69	1.02	0.00	0.80
50	5	2	75	0.43	0.95	0.40	0.80
50	5	2	116	0.00	1.12	0.36	1.03
50	5	3	151	0.35	0.89	0.00	0.95
50	5	3	205	0.00	0.42	0.25	0.93
50	6	1	17	0.00	0.32	0.22	0.89
50	6	1	39	0.07	0.34	0.22	0.85
50	6	2	79	0.36	0.66	0.34	1.09
50	6	2	117	0.33	0.61	0.24	1.09
50	6	3	152	0.00	0.46	0.30	0.98
50	6	3	213	0.00	0.47	0.00	1.20
50	7	1	20	0.00	0.43	0.00	0.81
50	7	1	44	0.00	0.00	0.00	0.73
50	7	2	82	0.79	0.54	0.00	0.84
50	7	2	118	0.80	0.71	0.00	0.74
50	7	3	155	0.39	0.33	1.02	0.80
50	7	3	224	0.52	0.56	0.45	0.86

false-negative and false-positive probabilities across laboratories. Visual assessment gave 100 % negative results (n=42) for 'blank' honeys and 100 % positive results (n=42) for test samples containing analyte at 100 % of the screening target concentration for all analytes. Hence, average false-positive and negative probabilities for visual assessment (at 100 % STC) are estimated to be, with 95 % confidence, *no higher than* 7 % for all analytes. This is estimated using a published approach [12], whereby *if 42 independent observations are made without error, then the underlying rate of errors may be 0 % and is,*

Table 3 Readsensor® results for honey fortified at 100 % of the screening target concentration

STC [%]	Lab	Day	ID	Sulfathiazole	Tylosin	Ciprofloxacin	Chloramphenicol
100	1	1	42	0.00	0.00	0.00	0.49
100	1	1	66	0.00	0.00	0.00	0.43
100	1	2	115	0.62	0.83	0.00	0.72
100	1	2	149	0.55	0.75	0.00	0.45
100	1	3	176	0.44	0.57	0.00	0.47
100	1	3	209	0.15	0.59	0.21	0.54
100	2	1	45	0.00	0.36	0.21	0.62
100	2	1	72	0.29	0.33	0.72	0.56
100	2	2	122	0.00	0.35	0.28	0.59
100	2	2	153	0.00	0.36	0.25	0.68
100	2	3	181	0.00	0.26	0.16	0.42
100	2	3	210	0.01	0.28	0.27	0.56
100	3	1	46	0.29	0.37	0.30	0.50
100	3	1	78	0.00	0.37	0.18	0.62
100	3	2	123	0.00	0.24	0.29	0.52
100	3	2	154	0.00	0.18	0.11	0.42
100	3	3	182	0.00	0.30	0.29	0.56
100	3	3	212	0.21	0.39	0.18	0.50
100	4	1	47	0.57	0.31	0.00	0.46
100	4	1	80	0.68	0.33	0.00	0.44
100	4	2	130	1.01	0.64	0.00	0.55
100	4	2	157	0.31	0.31	0.00	0.40
100	4	3	185	0.27	0.36	0.23	0.47
100	4	3	216	0.67	0.36	0.00	0.36
100	5	1	49	0.65	1.02	0.00	0.65
100	5	1	83	1.26	1.29	0.00	0.60
100	5	2	132	0.50	0.92	0.40	0.53
100	5	2	158	0.29	0.83	0.00	0.52
100	5	3	188	0.25	0.45	0.21	0.53
100	5	3	220	0.40	0.76	0.00	0.37
100	6	1	52	0.00	0.29	0.21	0.62
100	6	1	88	0.00	0.28	0.21	0.59
100	6	2	133	0.23	0.25	0.22	0.76
100	6	2	162	0.00	0.29	0.29	0.91
100	6	3	191	0.00	0.30	0.00	0.57
100	6	3	222	0.00	0.00	0.00	0.64
100	7	1	53	0.08	0.43	0.00	0.52
100	7	1	90	0.00	0.00	0.00	0.60
100	7	2	137	0.56	0.54	0.00	0.47
100	7	2	165	0.35	0.40	0.00	0.38
100	7	3	193	0.28	0.53	0.34	0.50
100	7	3	227	0.00	0.48	0.00	0.44

with 95 % confidence, less than $1-0.05^{1/42}=7$ %. One sample out of 42 containing chloramphenicol at 50 % of the target concentration was assessed to be negative. A more robust assessment of the method performance was conducted using the numerical data obtained using the Readsensor® (see below).

Instrumental assessment

Readsensor® results from the 42 measurements obtained over 3 days by the seven laboratories are shown in Tables 1, 2 and 3. These results were analysed (analysis of variance, random effects model with 'days' nested within laboratories) to produce one-tailed beta-expectation tolerance intervals [13, 14]. One-tailed beta-expectation tolerance intervals were used to provide the response ratio (a) above which 95 % of results would be expected to lie for samples that do not contain the analyte and (b) below which 95 % of results would be expected to lie for samples that contain the analyte at 50 and 100 % of the screening target concentration (Table 4). No results were excluded from the analysis.

Table 4 Analysis of Readsensor® results

STC [%]	Parameter	Sulfathiazole	Tylosin	Ciprofloxacin	Chloramphenicol
0	Mean	2.881	2.790	2.000	2.693
0	s_{lab}	0.629	0.551	0.263	0.365
0	s_{day}	0.154	0.394	0.172	0.158
0	s_r	0.370	0.301	0.204	0.268
0	*df*lab	6	6	6	6
0	*df*day	14	14	14	14
0	*dfr*	21	21	21	21
0	s_{mean}	0.247	0.230	0.111	0.148
0	s_R	0.746	0.741	0.375	0.480
0	*df*	11.45	17.28	20.91	16.29
0	t_{95}	1.78	1.73	1.72	1.74
0	Lower 0	1.48	1.44	1.33	1.82
50	Mean	0.317	0.482	0.208	0.799
50	s_{lab}	0.092	0.173	0	0.129
50	s_{day}	0.239	0.220	0.183	0.087
50	s_r	0.127	0.313	0.117	0.129
50	df_{lab}	6	6	6	6
50	df_{day}	14	14	14	14
50	df_r	21	21	21	21
50	s_{mean}	0.067	0.094	0.044	0.056
50	s_R	0.286	0.420	0.217	0.202
50	*df*	25.9	40.2	25.0	26.3
50	t_{95}	1.71	1.68	1.71	1.70
50	Upper 50	0.817	.206	0.587	1.156
100	Mean	0.260	0.427	0.132	0.536
100	s_{lab}	0.201	0.176	0.09	0.061
100	s_{day}	0.145	0.186	0.009	0.057
100	s_r	0.195	0.119	0.134	0.077
100	df_{lab}	6	6	6	6
100	df_{day}	14	14	14	14
100	df_r	21	21	21	21
100	s_{mean}	0.088	0.080	0.0399	0.029
100	s_R	0.315	0.282	0.162	0.114
100	*df*	26.6	24.9	26.0	35.1
100	t_{95}	1.70	1.71	1.71	1.69
100	Upper 100	0.817	0.928	0.416	0.734

Mean mean Readsensor® response across all results, s_{lab} estimate of between-laboratory component of reproducibility standard deviation, s_{day} estimate of between-day component of reproducibility standard deviation, s_r estimate of repeatability standard deviation, df_{lab} number of degrees of freedom associated with estimate of between-laboratory standard deviation, df_{day} number of degrees of freedom associated with estimate of between-day standard deviation, df_r number of degrees of freedom associated with the estimate of repeatability standard deviation, s_R estimate of reproducibility standard deviation, df_R number of degrees of freedom associated with the estimate of reproducibility standard deviation, t_{95} 95th quantile of the *t* distribution for df_R degrees of freedom, *Lower 0* lower end of a 95 % one-tailed beta-expectation tolerance interval for reader responses for honeys that do not contain residues, *Upper 50* upper end of a 95 % one-tailed beta-expectation tolerance interval for reader responses for honeys containing residues at 50 % of the target concentration, *Upper 100* upper end of a 95 % one-tailed beta-expectation tolerance interval for reader responses for honeys containing residues at 100 % of the target concentration

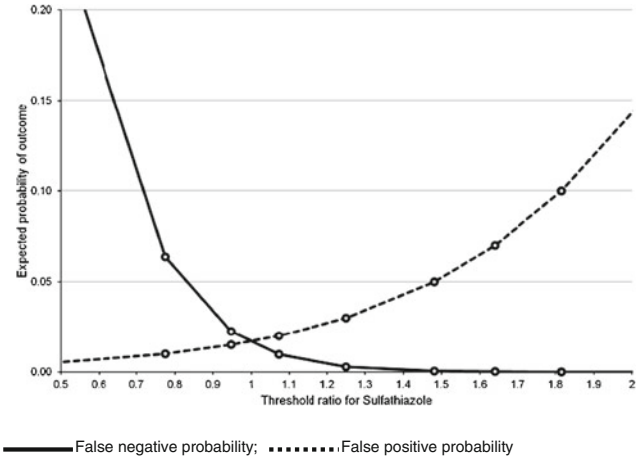

Fig. 2 Relation between expected false-positive probability, false-negative probability and the critical value for reader response for sulfathiazole

The intervals were estimated as follows. For each concentration, the size of the reproducibility standard deviation S_R was estimated by:

$$S_R = \sqrt{S_{\text{lab}}^2 + S_{\text{day}}^2 + S_r^2} \tag{1}$$

where $S_{\text{lab}}{}^2$ and $S_{\text{day}}{}^2$ were the between-lab and between-day components of variance, and $S_r{}^2$ was the error variance. The standard error associated with the mean (across laboratories and days) was estimated by:

$$s_{\text{mean}} = \sqrt{\frac{S_{\text{lab}}^2}{n_{\text{lab}}} + \frac{S_{\text{day}}^2}{n_{\text{day}}} + \frac{S_r^2}{n_r}} \tag{2}$$

where n_{lab} and n_{day} were the number of laboratories and days (seven labs; 21 days:3 days per laboratory) and n_r was the total number of measurements (42).

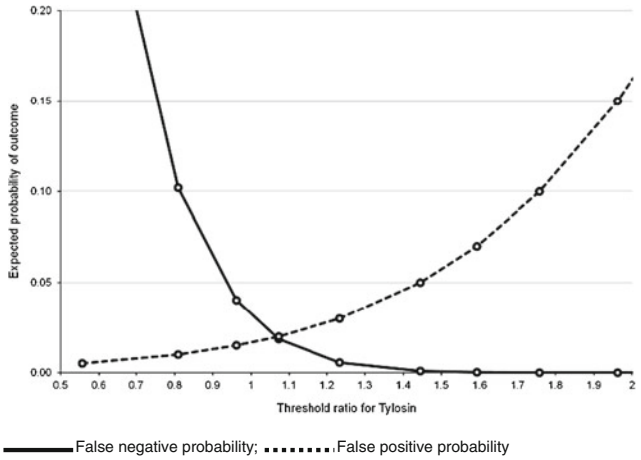

Fig. 3 Relation between expected false-positive probability, false-negative probability and the critical value for reader response for tylosin

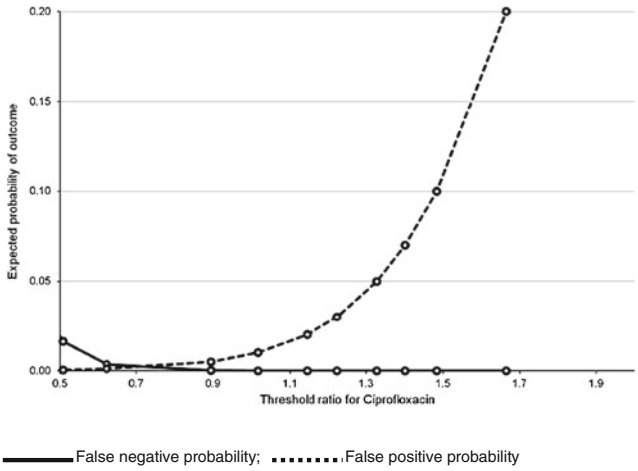

Fig. 4 Relation between expected false-positive probability, false-negative probability and the critical value for reader response for ciprofloxacin

Then, the value of the lower 95 % beta-expectation tolerance interval (y_L) was estimated by:

$$y_L = \overline{y} - t_{df,0.95}\sqrt{S_R^2 + S_{\text{mean}}^2} \tag{3}$$

For samples that did not contain the analyte, the value of the upper 95 % beta-expectation tolerance interval (y_u) was estimated by:

$$y_U = \overline{y} + t_{df,0.95}\sqrt{S_R^2 + S_{\text{mean}}^2} \tag{4}$$

for samples that contained the analyte at 50 and 100 % of the screening target concentration, where $t_{df,0.95}$ is the 95th percentile of the t distribution with df as degrees of freedom. The number of degrees of freedom, df_r, associated with the

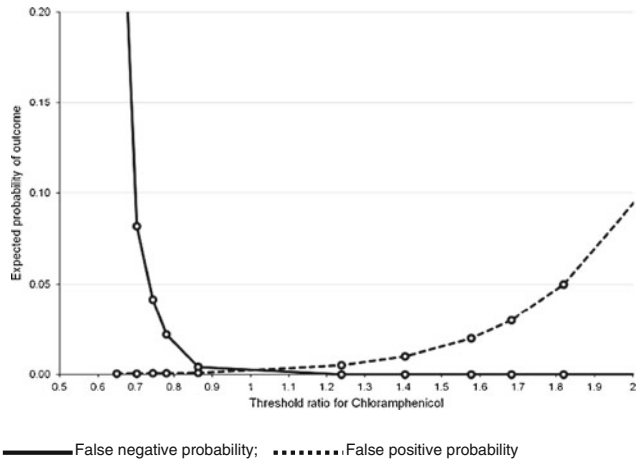

Fig. 5 Relation between expected false-positive probability, false-negative probability and the critical value for reader response for chloramphenicol

reproducibility standard deviation S_R, was estimated by the Welch–Satterthwaite equation:

$$df_R = \frac{\left(S_{\text{lab}}^2 + S_{\text{day}}^2 + S_r^2\right)^2}{\frac{S_{\text{lab}}^4}{df_{\text{lab}}} + \frac{S_{\text{day}}^4}{df_{\text{day}}} + \frac{S_r^4}{df_r}} \quad (5)$$

where df_{lab}, df_{day} and df_r are the number of degrees of freedom associated with each of the estimates of between-lab, between-day and residual error components of variation. Table 4 shows how beta-expectation tolerance intervals were calculated for each concentration of each analyte.

Beta-expectation tolerance intervals were inverted numerically to provide an estimate of the expected false-positive and false-negative probability associated with a range of different critical values for response ratios. This was done by adjusting the probability associated with the t value in Eqs. 3 and 4 until the value of Y_L (for false positives) or Y_U (for false negatives) matched the value of the critical level that was being assessed. The relation between the critical value that is chosen to make detection decisions and the expected false-positive and false-negative probabilities are shown in Figs. 2, 3, 4 and 5.

Data from Table 4 and Figs. 2, 3 and 4 clearly demonstrate that the instrumental (Readsensor®) measurement of the Bee4sensor® multiplex test is capable of detecting sulfathiazole and tylosin at or below a concentration of one half of the EU recommended concentrations [5], i.e. 25 and 10 μg kg^{-1}, respectively. In addition, ciprofloxacin could be reliably detected at 25 μg kg^{-1}. Furthermore, data previously obtained within our laboratory demonstrated that the method is also capable of detecting a wide range of sulfonamides (n=10) and fluoroquinolones (n=7) at similar concentrations (data not shown) to the 'representative analytes' that were selected for use in this inter-laboratory validation exercise. Unfortunately, the Bee4sensor® kit was not sufficiently sensitive to detect chloramphenicol at the EU MRPL of 0.3 μg kg^{-1} [4]. The method did, however, reliably detect this analyte at 5 μg kg^{-1} and would therefore be a valuable tool to give an early warning of honey containing higher concentrations. The detection of higher concentrations of chloramphenicol may be of more use to the honey industry when testing individual hives or smaller lots rather than a blended bulk product, as it is well known that high residue concentrations (between 200 and 20,000 μg kg^{-1}) can be found in honey taken directly from the hive for over 100 days after dosing [15]. It may also be of use to enforcement authorities in times of emergency when many samples need to be tested within a short timescale. In this case, the multi-sensor could be used to quickly identify honeys containing the highest concentration of this banned antimicrobial, before subjecting any suspect (or 'screen positive') sample to confirmatory analysis by LC-MS/MS [11].

Conclusions

The results produced by the dipstick, determined by visual observation or instrumental measurement of the four test lines versus the control line, were comparable.

The programmed set of ratios, to enable objective interpretation of the colour intensity with the software method, was deemed suitable to generate consistent Readsensor® results. The chosen critical values/threshold ratios were confirmed by the statistical evaluation of the data obtained from the analysis of blind samples at 0, 50 and 100 % of the EU recommended concentrations [5]. The inter-laboratory validation of the multisensor, with seven European laboratories, therefore demonstrated that the Bee4sensor® kit is both robust and capable of meeting the target concentrations set within the EU for sulfonamides (sulfathiazole), fluoroquinolones (ciprofloxacin) and tylosin. The method is also suitable for detecting chloramphenicol in honey at concentrations ≥5 μg kg^{-1}.

Acknowledgments The Con*ff*idence project is funded by the European Commission, call identifier FP7-KBBE-2007-1, grant agreement number 211326. The authors wish to thank the surveillance, LC-MS/MS teams and Michael Knaggs at Fera for analysing the test materials for residues of antibiotics and homogeneity. The authors thank the company Unisensor s.a. for supplying the test kits, instrumentation and their continuous support. The authors like to acknowledge the following laboratories for participating in this study: Fera (UK), CER (Belgium), Rikilt (The Netherlands), ILVO (The Netherlands), ANSES (France), Michaud (France) and one other.

References

1. Codex (2010) Discussion paper on veterinary drugs in honey production ftp://ftp.fao.org/codex/meetings/CCRVDF/ccrvdf19/rv19_10e.pdf. Accessed 3 Jan 2013
2. US Food and Drug Administration database. http://www.accessdata.fda.gov/scripts/animaldrugsatfda/details.cfm?dn=013-076. Accessed 3 Jan 2013
3. RASFF portal. https://webgate.ec.europa.eu/rasff-window/portal/#. Accessed 3 Jan 2013
4. European Commission (2003) Commission Decision 2003/181/EC of 13 March 2003 amending Decision 2002/657/EC as regards the setting of minimum required performance limits (MRPLs) for certain residues in food of animal origin. L71. 17–18
5. CRL guidance paper (2007) http://www.rivm.nl/bibliotheek/digitaaldepot/crlguidance2007.pdf. Accessed 3 Jan 2013
6. Community reference laboratories (2010). Guidelines for the validation of screening methods for residues of veterinary medicines (initial validation and transfer). http://ec.europa.eu/food/food/chemicalsafety/residues/Guideline_Validation_Screening_en.pdf. Accessed 3 Jan 2013
7. Reybroek W, Ooghe A, De Brabander H, Daeseleire E (2007) Validation of the tetrasensor honey test kit for the screening of tetracyclines in honey. J Agric Food Chem 55:8359–8366
8. Gaudin V, Hedou C, Verdon E (2013) Validation of two ELISA kits for the screening of tylosin and streptomycin in honey according to

the European Decision 2002/657/EC. Food Addit Contam Part A Chem Anal Control Expo Risk Assess 30(1):93–109
9. Conffidence (2010) Conffidence project website. http://www.conffidence.eu/img/enewsletter/enews_WP2b_antibiotics_newsletter_Nov_2010.pdf. Accessed 3 Jan 2013
10. Unisensor (2012) Bee4sensor® webpage. http://www.unisensor.be/en/catalog/antibiotics-28/bee4sensor-45.php. Accessed 3 Jan 2013
11. European Commission (2002) Commission decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. L221. 8–36
12. Brown LD, Cail TT, DasGupta A (2001) Interval estimation for a binomial proportion. Stat Sci 16(2):101–133
13. Mee RW (1984) β-expectation and β-content tolerance limits for balanced one-way ANOVA random model. Technometrics 26(3):251–253
14. Feinberg MJ (2007) Validation of analytical methods based on accuracy profiles. Chromatogr A 1158(1–2):174–183
15. Adams S, Heinrich K, Fussell R, Wilkins S, Thompson H, Ashwin H, Sharman M (2008) Study of the distribution and depletion of chloramphenicol residues in bee products extracted from treated honeybee (*Apis mellifera* L.) colonies. Apidologie 39:537–546