Quality Assurance and Safety of Crops & Foods, 2014; 6 (3): 299-307
SPECIAL ISSUE: *Food chemistry and safety*

Collaborative study for evaluating performances of a multiplex dipstick immunoassay for *Fusarium* mycotoxin screening in wheat and maize

V.M.T. Lattanzio[1], C. von Holst[2] and A. Visconti[1]

[1]*Institute of Sciences of Food Production, National Research Council of Italy, Via Amendola 122/O, 70126 Bari, Italy;*
[2]*European Commission, DG Joint Research Centre, Retieseweg 111, 2400 Geel, Belgium; veronica.lattanzio@ispa.cnr.it*

Received: 27 November 2013 / Accepted: 25 February 2014

RESEARCH ARTICLE

Abstract

An experimental protocol for evaluating performances of qualitative methods through interlaboratory validation has been designed and applied to a multiplex dipstick kit for the determination of major *Fusarium* toxins, namely zearalenone, T-2 and HT-2 toxins, deoxynivalenol, and fumonisins B_1 and B_2 in wheat and maize. The test is intended for screening of cereals for the presence/absence of these mycotoxins at maximum permitted levels established by European legislation. The response of the measurement is determined with a reader device. Decision on whether a sample is classified as positive or negative is taken by comparing the measured response with a cut off value which has been established based on the results from the method validation. The collaborative study involved 12 laboratories and included a training phase enabling the participants to familiarise with the protocol. The interlaboratory validation design consisted of three steps, namely (1) estimating the precision under reproducibility conditions; (2) establishing robust and laboratory independent cut off values for the dipstick response at target mycotoxin levels, assuming an acceptable rate of false negative results of 5%; and (3) assessment of the rate of false positive results of compliant samples. The total standard deviation of the response varied from 10 to 27% for the analyte/concentration/matrix combinations included in the study, indicating the assay ruggedness between different laboratories. The test resulted to be able to differentiate blank samples from samples contaminated at target mycotoxin levels with a false positive rate lower than 10% for zearalenone, deoxynivalenol and fumonisins, whereas unsatisfactory results were obtained for the sum of T-2 and HT-2 toxins. Finally, a critical evaluation of the outcomes of the whole validation study enabled to identify key issues for improving method transferability and reliability of the results.

Keywords: cereals, fumonisins, test kit validation, trichothecenes, zearalenone

1. Introduction

Rapid qualitative immunoassays enabling the detection of the presence/absence of a substance of interest at a target level are available for a variety of food contaminants. Immunoassay kits are commonly used for screening purposes to verify compliance with food and feed legislation requirements. Negative samples are considered as compliant whereas positive samples need to be re-analysed with confirmatory methods, reducing time and costs of the analysis. Therefore there is strong need for the development of such screening tests for the detection of residues and contaminants in food and feed matrices; this topic was addressed by the European FP7 funded project CON*ff*IDENCE (www.conffidence.eu).

Prior to their use under real world conditions the screening tests need to be validated against defined criteria. European legislation (EC, 2002) has established fitness for purpose criteria for screening tests when used within the frame of official control, placing emphasis on the rate of false negative results reflecting the probability that truly non-compliant samples were wrongly classified as negative. As specified in the Commission Decision the probability must not exceed 5% (β error). The rate of false positive results (α error) is also an important criterion, since negative samples

that are misclassified as positive need to be subjected to a superfluous confirmatory analysis.

For the detection of the major *Fusarium* mycotoxins, namely deoxynivalenol (DON), zearalenone (ZEA), T-2 (T2) and HT-2 toxins (HT2), and fumonisins B_1 and B_2 (FB_1, FB_2) in wheat and maize a specific multiplex lateral flow device has been developed within the EU project CON*ff*IDENCE (Lattanzio *et al.*, 2012). When applying this test, the response of the measurement is read by inserting the dipstick into a reader device. The numerical response is than compared with a cut off value and, depending on whether the result of analysis is below or above the cut-off value, the sample is considered as positive or negative. The method already passed single laboratory validation confirming that the screening test is fit for the intended purpose (Lattanzio *et al.*, 2013). A particular challenge of the validation step was that the currently available and internationally accepted validation guideline (Thompson *et al.*, 2002) did not fully addresses the specific validation requirements for qualitative screening tests. Therefore, an alternative validation protocol has been designed and applied to the single-laboratory validation of the screening test (Lattanzio *et al.*, 2013).

Following the successful validation results it was decided to subject the screening test to a collaborative validation study, where various laboratories applied the test on different samples. The purpose of the collaborative study was to assess the performance profile of the test under reproducibility conditions (ISO, 1994) and to establish a revised cut-off value based on the results reported by the participating laboratories. The present paper reports and discusses the results of the statistical assessment of this collaborative validation study.

2. Methods and materials

Materials

Methanol, and acetonitrile (both HPLC grade) were purchased from Mallinckrodt Baker (Milan, Italy). Ultrapure water was produced by a Millipore Milli-Q system (Millipore, Bedford, MA, USA). Standard mycotoxins (DON, ZEA, T2, HT2, FB_1, FB_2) were purchased from Fermentek (Jerusalem, Israel). Mycotoxin stock solutions were prepared by dissolving the commercial crystalline toxin in the appropriate solvent at concentration of 1 mg/ml. DON, T2, HT2 and ZEA were dissolved in acetonitrile, and FBs in acetonitrile:water (1:1). Multiplex dipstick kits (4mycosensor) and optical reader (Readsensor) were provided by Unisensor (Liège, Belgium). According to the provider instructions, dipstick kits were stored at 4 °C until use, and shipped in refrigerate conditions. In these conditions kits were stable for 12 months.

Multiplex dipstick immunoassay

Principle of the test

A basic scheme of the method is reported in Figure 1. In short, a ground maize sample is extracted sequentially with water then methanol. The ground wheat sample is extracted once with a mixture of methanol and water. The sample extract is diluted and directly analysed by the multiplex dipstick. The test kit employs a microwell of reagents containing four antibodies linked to gold particles and a dipstick made up of a nitrocellulose membrane where 4 specific capture lines (for ZEA, DON, T2+HT2, FB_1+FB_2, respectively) are located. In the presence of maize/wheat extract each antibody binds the corresponding mycotoxin before starting to run vertically on the dipstick in the direction of the capture lines. Within 10 min red lines appear from the background on the dipstick in correspondence of the capture lines. Results are interpreted by an optical reader measuring colour intensity, and calculating the ratio between each test line and the control line located on the top of the strip. Colour development at the control line ensures the validity of the test.

Sample preparation

For wheat, 10 g samples were weighed into a blender jar and extracted with 50 ml of a mixture of methanol:water (80:20, v/v) by blending at high speed for 2 min. The sample suspension was allowed to settle for 1 min. Then an aliquot of 100 µl of sample extract was added with 1,900 µl of running buffer (provided by the dipstick supplier) in an Eppendorf tube and gently mixed by hand. For maize, 10 g samples were weighed into a blender jar and extracted with 40 ml of water by blending at high speed for 2 min. Then 60 ml of methanol was added to the sample (without removing the first extract) and extracted again by high speed blending for 2 min. The sample suspension was allowed to settle for 1 min. Then an aliquot of 100 µl of sample extract was added to 900 µl of running buffer in an Eppendorf tube and gently mixed by hand.

Dipstick analysis

For dipstick analysis, 200 µl of diluted sample extract were transferred into the microwell and homogenised with the freeze-dried reagents by pipetting. The microwell was incubated for 10 min at 40 °C. Then the dipstick was added into the microwell and the sample was allowed to run for 10 min. After running, the filter was immediately removed by a spatula and the dipstick placed into the reader strip holder.

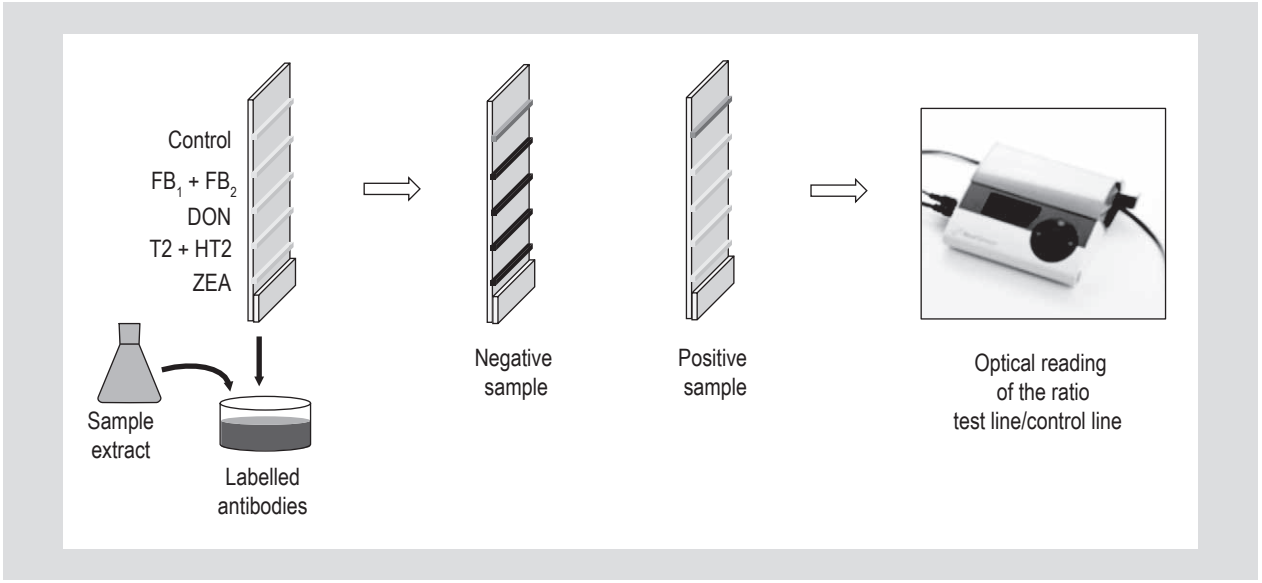

Figure 1. Scheme of dipstick immunoassay. The test kit employs a microwell of reagent containing four labeled antibodies and a dipstick where 4 specific capture lines (for zearalenone (ZEA), deoxynivalenol (DON), T-2 and HT-2 toxins (T2+HT2), and fumonisins B_1 and B_2 (FB_1+FB_2), respectively) are located. Results are interpreted by the optical reader measuring the ratio between each test line and the control line.

Organisation of collaborative study

Twelve laboratories (listed in Table 1) from 5 different European countries, representing a cross-section of government, food control, and food industry affiliations were involved in the collaborative trial. Each laboratory had to analyse duplicate samples of blank samples and samples with the analyte at target level.

Prior to the final validation study all laboratories had to pass a training phase, in which they analysed various samples.

Table 1. Participating laboratories.

Participant	Country
Barilla G.R. F.lli SpA	Italy
Centre d'Economie Rurale, C.E.R. Groupe	Belgium
DG Joint Research Centre, Institute for Reference Materials and Measurements	Belgium
Euroclone S.P.A.	Italy
Food and Environment Research Agency	UK
Ghent University	Belgium
Institute of Sciences of Food Productions, National Research Council of Italy	Italy
MasterLab, Nutreco	the Netherlands
RIKILT – Institute of Food Safety	the Netherlands
Unisensor	Belgium
University of Natural Resources and Applied Life Sciences (IFA-Tulln)	Austria
Wal.Agri SA	Belgium

The purpose of the training phase was to familiarise the laboratories with the correct execution of the method protocol. Each participant received:

1. For the training and final validation study two identical sample sets. Each sample set contained 8 samples of 10 g, which were blind duplicates of (1) the blank wheat and maize materials (negative samples); and (2) the wheat and maize materials with the analytes at target level (positive samples). The individual wheat and maize samples were labelled with the codes W1,W2, W3 and W4 and M1, M2, M3 and M4, respectively.
2. One dipstick kit containing 16 dipsticks, 16 reagent wells, and a batch of running buffer.
3. Microwell incubator.
4. Strip reader, and instructions for reader installation and calibration.
5. The protocol of the method. For the final validation study the revised version of the method protocol in standard operating procedure format was sent to the participating laboratories.
6. Forms for reporting the analytical data to the organiser of the study.
7. Material receipt form.

Statistics

The reported results from the laboratories were statistically evaluated by applying the same validation approach that was already used in the single laboratory validation of the test. Details of the procedure are given in the corresponding paper (Lattanzio *et al.*, 2013). The validation procedure consisted of the following three steps:

1. The average values of the numerical responses reported by the laboratories and the corresponding precision under repeatability and reproducibility conditions were estimated separately for the blank and positive samples by means of analysis of variance (ANOVA) (ISO, 1994). We applied robust statistics for the ANOVA (Analytical Methods Committee, 1998), to address adverse effects of extreme values on the calculated precision estimates. In short, robust statistics does not require the identification and elimination of outliers from ANOVA, just extreme values are truncated. The precision was expressed in terms of relative standard deviation of repeatability (RSD_r) and relative standard deviation of reproducibility (RSD_R). The calculation was done using an Excel macro available from the Analytical Methods Committee (http://www.rsc.org/membership/networking/interestgroups/analytical/amc/software/RobustStatistics.asp).
2. The cut off value was calculated using the average value and the standard deviation of reproducibility calculated in the previous step of the results from the positive samples ($SD_{positive\ samples}$) according to the following equation:

$$\text{Cut off value} = \text{average value} + \text{t-value}_{(\text{one-sided, } \beta=0.05)} \times SD_{positive\ samples} \quad (1)$$

3. Finally we used the results from the negative samples, to estimate the rate of false positive results. This was achieved by calculating the t-value as specified here:

$$\text{t-value} = \frac{(\text{average}_{negative\ samples} - \text{cutv})}{SD_{negative\ samples}} \quad (2)$$

Where cutv is the cut off value and $SD_{negative\ samples}$ is the standard deviation of reproducibility of the negative samples.

Preparation of test materials

The following test materials were prepared for the collaborative trial: blank samples and samples containing the mycotoxins at the target levels, i.e. 100 µg/kg ZEA, 250 µg/kg T-2 and HT-2 each, and 1,750 µg/kg DON in wheat, and 350 µg/kg ZEA, 250 µg/kg T-2 and HT-2 each, 1,750 µg/kg DON, 3,000 µg/kg FB_1, and 1000 µg/kg FB_2 in maize.

Two batches of 1 kg each of blank wheat and two batches of 1 kg each of blank maize were selected for preparation of materials for training and test phase of the study. All cereal samples were finely ground with a Tecator Cyclotech 1093 (International PBI, Milan, Italy) laboratory mill equipped with a 500 µm sieve.

Prior to the use of the blank materials for the collaborative study, the absence of target analytes was checked by applying different reference methods specific for a mycotoxin or group of mycotoxins, namely for FB_1 and FB_2 (Sydenham *et al.*, 1996), for T2 and HT2 (Pascale *et al.*, 2012), for DON (MacDonald *et al.*, 2005b), and for ZEA (MacDonald *et al.*, 2005a). Relevant detection limits were 10 µg/kg for FB_1 and FB_2, 8 µg/kg for T2 and HT2, 20 µg/kg for DON, and 3 µg/kg for ZEA. Concentrations of the target mycotoxins in the selected uncontaminated wheat and maize materials were below the relevant detection limit.

The test samples containing the mycotoxins at target level were prepared by fortifying blank samples with mixed solutions of mycotoxins. In detail, the solutions were freshly prepared by drying down appropriate amounts of stock solutions of each mycotoxin and redissolving them in 50 ml acetonitrile. Two different solutions were prepared, namely solution 1 for wheat containing 1 µg/ml ZEA, 2.5 µg/ml T2, 2.5 µg/ml HT2, and 17.5 µg/ml DON, and solution 2 for maize containing 3.5 µg/ml ZEA, 2.5 µg/ml T2, 2.5 µg/ml HT2, 17.5 µg/ml DON, 30 µg/ml FB_1, and 10 µg/ml FB_2. Then 500 g of wheat and maize were spiked with 50 ml of solution 1 and 2, respectively. After the spiking step, the materials were left overnight at room temperature to allow solvent evaporation and homogenised by passing them again through the laboratory mill equipped with a 500 µm sieve. Finally subsamples of 10 g of the blanks and spiked materials were taken, filled in labelled plastic boxes, sealed and stored at -20 °C until dispatch.

3. Results and discussion

Training phase

Prior to the final validation study all laboratories had to pass a training phase, in which they analysed blank and contaminated wheat and maize samples. The purpose of the training phase was to familiarise the laboratories with the correct execution of the method protocol. In addition the laboratories gave quite useful comments, especially regarding the description of the method protocol. Based either on these comments or on the results of the training phase, the method protocol was slightly revised, in order to make it sharper and to minimise deviations from the standard procedure.

Main outcome of the training phase was the report by different participants of unexpected low intensity values for the control line (i.e. 30% of the expected value) in some samples. These low values of the control line gave extremely high values of the test line/control line ratios, causing misclassification of the samples (false negatives). To overcome this problem, the method protocol was modified by specifying a threshold for the control line intensity ensuring the validity of the test. The participants were asked to repeat the test if the control line intensity resulted to be lower than this threshold.

Collaborative study

The numerical responses of the analysis of the test samples as reported by 12 laboratories are shown in Table 2-5. In total, 24 values for each analyte/concentration/matrix combination were available and used for statistical assessment. Since robust statistics have been applied to estimate the average value across all results and the corresponding performance profile, all the results obtained

Table 2. Measured dipstick response (test line/control line ratio) for zearalenone reported by the individual laboratories obtained with the different test materials.

Lab no.	Wheat				Maize			
	Blank		Target level		Blank		Target level	
	a	b	a	b	a	b	a	b
1	2.28	2.13	1.41	2.19	3.77	3.64	1.23	2.24
2	1.94	2.61	1.12	1.15	2.71	2.70	1.08	1.03
3	2.09	1.99	1.31	1.53	7.76	2.37	1.17	1.17
4	1.97	1.87	1.18	1.21	2.06	1.96	1.27	1.15
5	2.38	2.24	1.40	1.72	2.75	2.92	1.28	1.42
6	2.46	2.25	1.37	1.44	2.87	2.74	1.24	1.19
7	1.87	2.20	1.19	1.27	2.48	2.10	1.09	0.92
8	2.45	1.65	1.09	1.09	1.86	1.96	1.02	1.05
9	2.29	2.26	1.43	1.80	2.74	3.11	1.34	1.26
10	2.00	1.95	1.19	1.42	2.19	2.30	0.87	0.98
11	2.65	2.46	1.56	1.45	3.37	3.49	1.30	1.49
12	1.93	1.63	1.27	1.21	2.13	2.21	1.18	1.23

a, b = replicate analysis of the same sample

Table 3. Measured dipstick response (test line/control line ratio) for T-2 and HT-2 toxins reported by the individual laboratories obtained with the different test materials.

Lab no.	Wheat				Maize			
	Blank		Target level		Blank		Target level	
	a	b	a	b	a	b	a	b
1	2.24	2.28	1.92	2.74	2.74	2.37	2.19	2.41
2	1.94	2.89	1.63	1.87	1.87	3.18	3.05	2.27
3	2.34	2.27	1.95	2.23	9.76	2.75	2.49	2.55
4	2.13	2.17	1.63	1.59	2.19	2.20	2.41	2.24
5	2.56	2.61	2.07	2.45	3.04	3.27	2.63	2.75
6	2.68	2.48	2.08	2.12	3.14	3.04	2.59	2.55
7	2.02	2.45	1.80	1.84	2.59	2.30	2.19	1.61
8	2.07	1.91	1.69	1.61	2.16	2.17	2.16	2.12
9	2.50	2.53	2.01	2.37	3.08	3.70	2.75	2.46
10	2.37	2.25	1.79	2.02	2.46	2.56	1.87	2.12
11	2.86	2.65	2.34	2.26	3.78	3.89	2.84	3.14
12	2.08	1.67	1.73	1.69	2.34	2.36	2.26	2.36

a, b = replicate analysis of the same sample.

Table 4. Measured dipstick response (test line/control line ratio) for deoxynivalenol reported by the individual laboratories obtained with the different test materials.

Lab no.	Wheat				Maize			
	Blank		Target level		Blank		Target level	
	a	b	a	b	a	b	a	b
1	3.48	3.45	1.24	1.45	3.37	2.64	1.23	1.17
2	3.82	4.55	0.78	0.87	3.04	3.46	0.97	1.11
3	3.36	3.30	1.08	1.03	9.55	3.14	1.19	1.16
4	3.18	2.99	0.95	0.94	2.68	2.47	1.26	1.18
5	3.47	3.82	1.01	1.16	3.34	3.42	1.09	1.25
6	3.97	3.64	0.96	0.98	3.16	3.15	1.10	1.15
7	3.05	3.42	0.98	1.00	2.99	2.54	1.01	0.83
8	2.89	2.61	1.08	1.04	2.48	2.6	1.30	1.31
9	3.13	3.96	1.01	1.09	3.47	3.77	1.22	1.13
10	3.42	3.46	0.94	1.08	3.05	2.78	1.00	1.08
11	4.3	4.22	1.09	1.22	4.00	4.12	1.11	1.16
12	3.22	2.82	0.99	1.04	2.71	2.99	1.32	1.28

a, b = replicate analysis of the same sample.

from participant laboratories were considered for the statistical analysis, without outlier exclusion.

Table 5. Measured dipstick response (test line/control line ratio) for fumonisins B_1 and B_2 reported by the individual laboratories obtained with the different test materials.

Lab no.	Maize			
	Blank		Target level	
	a	b	a	b
1	3.06	2.43	0.41	0.50
2	3.07	3.51	0.38	0.36
3	3.21	3.17	0.45	0.31
4	2.47	2.50	0.47	0.40
5	3.05	3.34	0.47	0.57
6	3.16	3.21	0.41	0.34
7	2.97	2.44	0.43	0.37
8	2.18	2.34	0.56	0.50
9	3.19	3.65	0.56	0.50
10	2.92	2.83	0.42	0.44
11	4.04	3.99	0.43	0.44
12	2.57	2.64	0.48	0.45

a, b = replicate analysis of the same sample.

The results of the statistical assessment in terms of the precision profile and the calculated cut off value are shown in Table 6. The RSD_r values ranged from 5.5 to 12% and in some cases were lower than the corresponding values obtained in the single laboratory validation (Lattanzio *et al.*, 2013) which ranged from 8 to 16%. Furthermore the RSD_R values ranged from 10 to 27%, which was considered acceptable for the type of method evaluated in this paper.

Quite different results were obtained when calculating the false positive rate (Table 6) resulting very low for DON and fumonisins, and higher but still acceptable for ZEA (below 10%). Unacceptable rate of false positive results was obtained for the sum of T2+HT2 either in wheat (61%) and maize (76%).

Figure 2 shows, for two examples obtained with maize, the numerical responses reported by the laboratories along with the calculated cut-off value. For DON in maize (Figure 2A) the plot clearly indicated that the results from the blank samples and the samples fortified with the analyte at target level formed two well defined groups that were easily separated by the cut off value. This aspect was also mirrored by the very low rate of false positive samples, which was 0.2% (Table 7). In contrast to DON, very poor results were obtained for T2+HT2 (Figure 2B), since numerical responses from the positive and negative samples strongly overlapped, thereby leading to a very high rate of false positive results of 76% (Table 7).

Table 6. Precision profile of each analyte/matrix/concentration combination and cut off values.

		Wheat			Maize			
***Fusarium* toxin**		**ZEA**	**T2+HT2**	**DON**	**ZEA**	**T2+HT2**	**DON**	**FB_1+FB_2**
Blank	Mean response (T/C ratio)[1]	2.1	2.3	3.5	2.7	2.8	3.1	3.0
	RSD_R (%)[2]	12	13	15	27	20	18	16
	RSDr (%)***	8.8	6.9	7.5	5.5	8.6	8.5	7.9
Target level	Mean response (T/C ratio)[1]	1.4	2.0	1.0	1.2	2.4	1.2	0.4
	RSD_R (%)[2]	18	17	10	17	20	10	17
	RSDr (%)[3]	11	9.3	7.3	7.9	7.8	6.3	12
Cut off value (T/C ratio[4])		1.8	2.5	1.2	1.5	3.0	1.4	0.6

[1] The mean values of the response (ratio test line/control line; T/C ratio) are calculated from the results of 24 experiments.

[2] RSD_R (%) = relative standard deviation of reproducibility obtained from the collaborative study.

[3] RSDr (%) = relative standard deviation of repeatability obtained in the previous in-house validation study (Lattanzio *et al.* 2013).

[4] T/C ratio = ratio test line/control line.

DON = deoxynivalenol; FB_1 = fumonisin B_1; FB_2 = fumonisin B_2; HT2 = HT-2 toxin; T2 = T-2 toxin; ZEA = zearalenone.

Comparison of the performance profile of the screening test obtained in the single laboratory validation (Lattanzio *et al.*, 2013) and in the collaborative trial presented in this paper clearly demonstrated some significant differences between the two studies. When focusing on the rate of false positive results for the blank samples, comparable results were obtained for DON in both matrices, which were even better for FB_1+FB_2 in the collaborative study. For ZEA in maize and wheat the method performed worse in this study, since the rates of false positive results were 9.6 and 5.7%, compared to 0.85 and 0.58%, respectively, from the single-laboratory validation (Lattanzio *et al.*, 2013). However, being lower than 10%, the rate of false positives for ZEA could be considered still acceptable. In contrast with the present study, excellent values were obtained for T2+HT2, since the corresponding values from the single laboratory validation were 0.3 and 4.1%. When looking at possible reasons for such a huge difference, two factors have to be considered, namely (1) the average values across all laboratories for the negative and positive samples; and (2) the values for the standard deviations used for the calculation of the cut off values. For the single-laboratory validation the intermediate precision was used, whereas in the current study the standard deviation of reproducibility was applied. The principle applies that a high difference of these average values and low standard deviations favour

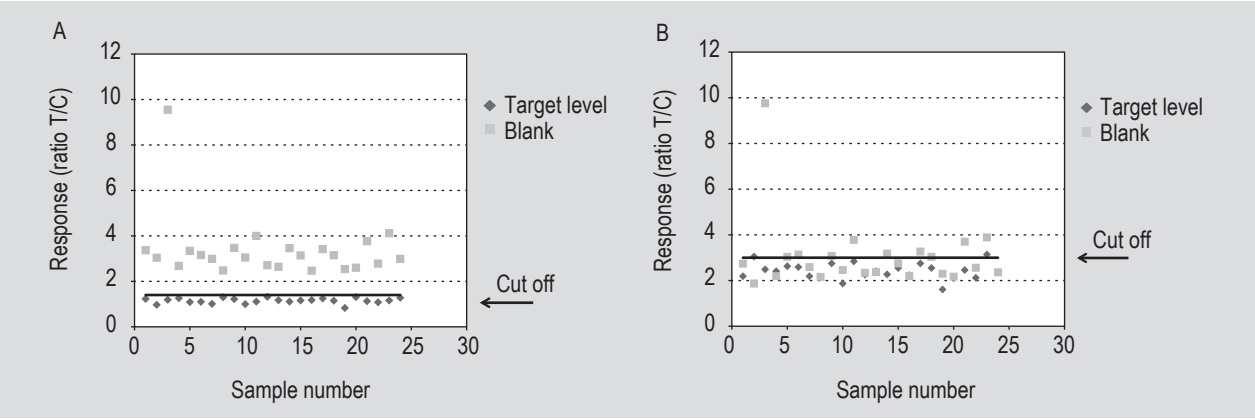

Figure 2. Dipstick response for (A) deoxynivalenol (DON) and (B) T-2 and HT-2 toxins (T2+HT2) in maize: examples showing the good (DON) and bad (T2+HT2) performances for the analysis of blank samples and samples fortified at target mycotoxin levels. The cut off values were calculated from the results obtained with the analyte at target level, assuming 95% of the samples correctly classified as positive. T/C = ratio test line/control line.

Table 7. Rate of false positive results (%) for blank wheat and maize samples.

	Fusarium toxin	Percentage false positives (%)
Wheat	ZEA	9.6
	T2+HT2	76.0
	DON	0.02
Maize	ZEA	5.7
	T2+HT2	61.0
	DON	0.2
	FB_1+FB_2	<0.01

DON = deoxynivalenol; FB_1 = fumonisin B_1; FB_2 = fumonisin B_2; HT2 = HT-2 toxin; T2 = T-2 toxin; ZEA = zearalenone.

a good performance of the test. For the specific case of T2+HT2 in both matrices, the precision of single laboratory validation varied from 10 to 17%, which was lower than the precision of the collaborative study (ranging from 13 to 20%). However, since this difference of precision values was minor, we concluded that the precision profile was not responsible for the bad performance of the method. When comparing the average values of the results from test samples, similar values were obtained in both validation studies for the blanks, since the ratios test line/control line ranged between 2.1 and 2.8. In contrast, in the single-laboratory validation the average of the results of positive samples dropped as expected to 1 to 1.3, whereas the corresponding values in the collaborative study remained at 2 to 2.3, which were very close to those of the blank samples. In other words, in the collaborative study, the presence of T2 and HT2 in samples contaminated at target levels gave a poor signal inhibition (i.e. inhibition of colour development) at the test line. This might be ascribed to a low affinity of the anti-T2 antibody, unable to bind a sufficient amount of T2/HT2 in the contaminated cereal extract in the 10 min of incubation time. Provided that the efficiency of the applied extraction procedure has been demonstrated in a previous study (Lattanzio *et al.*, 2012) these results could also be attributed to unexpected performance variations between different kit lots (buffer, strip coating, reagent lyophilisation). Therefore, one important outcome of the present study is that variation of kit performances between different production lots should be included in the in-house validation design.

The overall results confirm the general applicability of the test to separate samples with mycotoxins at the target levels (EU maximum permitted levels) from blank samples for ZEA, DON and fumonisins. However the kit still needs to be improved in order to properly discriminate blank samples from samples contaminated with T2/HT2 toxin, also in view of the recommended levels recently issued by the EC (EC, 2013).

4. Conclusions

In the present work an experimental design has been proposed for evaluating performance characteristics of qualitative immunoassays through interlaboratory validation. The applicability of the developed design has been demonstrated for a qualitative screening immunoassay based on the use of a multiplex dipstick kit for the determination of major *Fusarium* toxins considered by EU regulation. Matrices chosen for method validation were raw wheat and maize, representing the major field of application for the kit under study.

The validation study, involving 12 laboratories, delivered the precision profile of the method under reproducibility conditions and demonstrated method ruggedness among the different participant laboratories. Evaluation of false positive rates revealed that the test was able to differentiate blank samples from samples contaminated at target mycotoxin levels with a false positive rate lower than 10% for ZEA, DON and fumonisins, whereas unacceptable results were obtained for the sum of T2+HT2. Finally, a critical evaluation of the outcomes of the whole validation study enabled to identify key issues for improving method transferability and reliability of the results.

Acknowledgements

The research conducted within the Con*ffi*dence project and leading to these results has received funding from the European Commission's Seventh Framework Programme (call identifier FP7-KBBE-2007-1) under grant agreement no. 211326. We thank the laboratories listed in Table 1 for participation in the collaborative study. We are also very grateful to Unisensor (Liège, Belgium) for providing the 4-mycosensor kit.

References

Analytical Methods Committee, 1989. Robust statistics – how not to reject outliers. Part 1. Basic concepts. Analyst 114: 1693-1697.

European Commission (EC), 2002. Commission Decision 2002/657/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. Official Journal of European Union L22: 1-8.

European Commission (EC), 2013. Commission Recommendation 2013/165/EC of 27 March 2013 on the presence of T-2 and HT-2 toxin in cereals and cereal products. Official Journal of European Union L9: 1-12.

International Standardisation Organisation (ISO), 1994. ISO 5725-3:1994(E). Accuracy (trueness and precision) of measurement methods and results – part 3: intermediate measures of the precision of a standard measurement method. ISO, Geneva, Switzerland.

Lattanzio, V.M.T., Nivarlet, N., Lippolis, V., Della Gatta, S., Huet, A.C., Delahaut, P., Granier, B. and Visconti, A., 2012. Multiplex dipstick immunoassay for semi-quantitative determination of *Fusarium* mycotoxins in cereals. Analytica Chimica Acta 718: 99-108.

Lattanzio, V.M.T., Von Holst, C. and Visconti, A., 2013. Experimental design for in house validation of a screening immunoassay kit. The case of a multiplex dipstick for *Fusarium* mycotoxins in cereals. Analytical and Bioanalytical Chemistry 405: 7773-7782.

MacDonald, S.J., Anderson, S., Brereton, P., Wood, R. and Damant, A., 2005a. Determination of zearalenone in barley, maize and wheat flour, polenta, and maize-based baby food by immunoaffinity column cleanup with liquid chromatography: interlaboratory study. Journal of AOAC International 88: 1733-1740.

MacDonald, S.J., Chan, D., Brereton, P., Damant, A. and Wood, R., 2005b. Determination of deoxynivalenol in cereals and cereal products by immunoaffinity column cleanup with liquid chromatography: interlaboratory study. Journal of AOAC International 88: 1197-1204.

Pascale, M., Panzarini, G. and Visconti, A., 2012. Determination of HT-2 and T-2 toxins in oats and wheat by ultra-performance liquid chromatography with photodiode array detection. Talanta 89: 231-236.

Sydenham, E.W., Shephard, G.S., Thiel, P.G., Stockenström, S., Snijman, P.W. and Van Schalkwyk, D.J., 1996. Liquid chromatographic determination of fumonisins B1, B2, and B3 in corn: AOAC-IUPAC collaborative study. Journal of AOAC International 79: 688-696.

Thompson, M., Ellison, S.L.R. and Wood, R., 2002. Harmonized guidelines for single laboratory validation of methods of analysis. Pure and Applied Chemistry 74: 835-855.