Initial validation of FoodSmartphone screening methods

Introduction

This document is a short manual for the absolute minimum effort required by the FoodSmartphone project with respect to the initial in-house validation of newly developed screening methods. This minimum is based on what reviewers would normally require for peer-reviewed publication of your method, on current EU regulations[1,2] and opinions of EU reference laboratories[3] with respect to the validation of screening methods. Herein we classify your new FoodSmartphone method by detection principle as a biochemical method that detects molecular interactions. With respect to the degree of quantitation we classify your new method as (i) a qualitative screening method, just providing a yes/no response above or below a cut-off level and/or as (ii) a semi-quantitative screening method, i.e. the same but including a numerical readout and a calibration curve (provided by a smartphone). The result of such a screening method is either "negative" (sample is compliant versus a regulatory limit) or "suspect" (sample must be send to a lab for confirmatory analysis providing unambiguous identification and quantification). The most important performance characteristic for a screening method is the false compliant rate or β error which must be (equal or) lower than 5%. Following this classification and logical order of events your new method can and will never compete with mass spectrometric confirmatory methods: your benchmark method to beat is a conventional screening method such as ELISA, biosensors, microbiological inhibition tests, etc.

How to get started and what to do

- You need a kind of a manual for your newly developed FoodSmartphone screening method (we will call that a standard operating procedure;-).
- Your need standard solutions of your target substance(s).
- You must know the EU regulatory limit (MRL/ML/MRPL/RPA) for your target substance in your target sample matrix. Also knowing regulatory limits outside the EU would be quite useful (in some cases there are even no limits in the EU but regulatory guidance does exist from outside the EU).
- You must define the screening target concentration (STC), which should be set at or below the regulatory limit, preferably be set at one half of the regulatory limit.
- You need 20 different blank sample matrices (truly different, so for example in the case of milk: not 20 semi-skimmed milk samples from one batch of milk, but milks having different levels of fat, pasteurized and UHT-treated, from different brands and different lots!).
- The same 20 different blank sample matrices but fortified with your target substance at the screening target concentration (STC). Preferably the spiking level is verified by a quantitative instrumental analysis method.
- In case of a (ii) semi-quantitative screening method using a calibration curve: 5 additional fortified samples spiked around the STC if a matrix calibration curve is required. If you have experimental evidence that the sample matrix will not influence your screening method at all then the calibration curve may be constructed in assay buffer.
- Perform analyses of the 20 blank and the 20 fortified samples (plus 5 in case of a calibration curve method) with your new FoodSmartphone method, preferably divided over five different days and preferably in random blind order. Record the result as (i) yes/no response and/or (ii) the numerical value obtained.

Data evaluation for (i) qualitative screening methods

1. At least 19 out of the 20 spiked samples should have been correctly classified as "suspect", i.e. provided a result different from the blank samples. In that case the false compliant rate is equal or lower than 5% and the CCβ value is less than or equal to the STC (equals a kind of cut-off value in this situation). If your new method did not achieve this then it is simply not sensitive enough (not fit-for-purpose) and should be adjusted followed by a new initial validation.

---

[1] Commission Regulation (EU) No 519/2014 of 16 May 2014.
[2] Commission Decision 2002/657/EC.
[3] Community Reference Laboratories Residues, Guidelines for the validation of screening methods for residues of veterinary medicines, 20 January 2010.

2. Also assess how many of the 20 blank samples were incorrectly classified as "suspect" and document this false-positive rate. If your method has a high false-positive rate then nobody will use your method for screening since too many samples have to be unnecessarily send to the lab for confirmatory analysis.....

Data evaluation for (ii) semi-quantitative screening methods

1. Construct a calibration curve having at least six concentration levels (including zero), describe the working range of the curve, the mathematical equation end the goodness-of-fit.
2. Determine the numerical values for the 20 plus 20 samples analysed.
3. Determine the cut-off levels and CCβ according to procedure A or B below:

Procedure A (visualised in Annex I):
The numerical range of responses in the blank samples is examined and the highest blank response is noted [the lowest in case of a competitive inhibition immunoassay format]. Next the numerical range of responses in the spiked samples is examined and the lowest response of the spiked samples is noted [the highest in case of a competitive inhibition binding assay format]. If there is no overlap between the ranges of the blanks and the spiked samples then the CCβ level of the screening method is less than or equal to the STC and the false compliant rate lower than 5%. If your new method did not achieve this then it is simply not sensitive enough (not fit-for-purpose) and should be adjusted followed by a new initial validation.

Procedure B (visualised in Annex II):
The mean response (B) and standard deviation (SD_B) of the numerical values of the blank samples is calculated. The mean response (M) and standard deviation (SD_M) of the numerical values of the spiked samples is calculated. Next the threshold value (T) and the cut-off factor (Fm) is determined using the equations $T = B + 1.64 \times SD_B$ and $Fm = M - 1.64 \times SD_M$ [in case of a competitive inhibition binding assay format these equations should be $T = B - 1.64 \times SD_B$ and $Fm = M + 1.64 \times SD_M$].
If Fm > T [in case of a competitive inhibition binding assay format: Fm < T] then the CCβ is lower than the STC and the false compliant rate is lower than 5% and also the false positive rate is lower than 5%.

What's to be done in addition

1. Specificity: since your method relies on biorecognition the specificity should be documented by assessment of the cross-reactivity of related regulated substances and unrelated (regulated) substances that may occur in the sample matrix at similar or much higher levels. The cross-reactivity profile may be measured by standard solutions in assay buffer but the most relevant results from those experiments should be verified in spiked sample matrix as well.
2. Benchmarking: we hope that your newly developed FoodSmartphone screening method is not only simpler to use and cheaper, but also performing equal or better than existing screening methods. So the performance characteristics of your new method should be benchmarked against (at least) one conventional established screening method for your target substance in your target matrix.

The final proof of the FoodSmartphone concept

The final goal of the project is to have new on-site screening methods that can be operated by non-experts outside the lab. So go out there and simply demonstrate it! Use for example students or volunteers or friends or family members or farmers or citizens to operate your newly developed method and record their comments and suggestions.

Publication of FoodSmartphone applications

All peer reviewed publications of FoodSmartphone applications must meet the minimum standards of validation outlined in this document. Prior to manuscript submission it is the responsibility of the corresponding author to check and approve the validation data.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 720325.

Annex I. Data evaluation for semi-quantitative screening methods according to procedure A

Example A:

- MRL = 1.0 µg/kg
- Desired Screening Target Concentration = 0.5 µg/kg

Twenty (or multiples thereof) different blank matrix samples are selected. Replicates of these samples are spiked at the Screening Target Concentration, in this case 0.5 µg/kg.

The matrix blank samples and spiked samples are analysed, preferably over a number of different days. The range of responses in the blank samples is examined. The **highest response in the blank** samples is noted – in this case it is 0.137 units. The **lowest response in the spiked** samples is noted – in this case it is 0.252 units.

In the case shown none of the responses of the spiked samples overlaps with the range of responses of the blanks. Therefore we can say that **the CCβ of this screening method is less than or equal to 0.5 µg/kg**.

In the example shown we can see that the lowest response is 0.252. Therefore **the Cut-Off Level of this test is 0.252 units.** Any sample giving a response greater than this level is deemed to be a 'screen positive' and exceeds the CCβ of the screening method.

Sample Number	Negative Samples	Spike @ 0.5 µg/kg
1	0.000	0.355
2	0.090	**0.252**
3	0.000	0.532
4	0.000	0.554
5	0.000	0.408
6	0.070	0.501
7	0.000	0.524
8	0.015	0.559
9	0.000	0.471
10	0.010	0.661
11	0.070	0.642
12	0.129	0.724
13	0.046	0.596
14	0.034	0.599
15	0.041	0.640
16	**0.137**	0.750
17	0.112	0.655
18	0.120	0.660
19	0.132	0.695
20	0.063	0.635

For this test, as a batch acceptability criterion, the response generated by the "Screen Positive Control Sample" must be ≥ 0.25 units otherwise the batch is rejected.

Annex II. Data evaluation for semi-quantitative screening methods according to procedure B

Graphical representation of threshold value T and "cut-off" factor Fm.