

## **NMKL Protocol No. 7, 2014:**

### **Requirement to internal (inhouse / single laboratory) validation of presumptive NMKL microbiological methods**

#### **1. Introduction**

It is important to document the performance of a method. Even if methods are not validated in collaborative (interlaboratory) studies, there should be some validation data made available providing information about the applicability of the method.

#### **2. Work flow in NMKL (for non-collaborative validated method)**

- The referee elaborate a draft method in cooperation with appointed contact persons
- The laboratory of the referee, or another laboratory assigned by the referee, carries out an internal validation of the method. Results from PT schemes and results from internal validation in establishment of measurement uncertainty can be used.
- The results are reviewed by the referee and the contact persons
- The results are included in the draft method
- The draft method, including the results, is reviewed by the national committees, first by the national committee of the referee.
- The chair of Sub-committee 2 and the Secretary General perform the final review of the method
- The method is reviewed for revision after 5 years, or whenever needed

#### **3. Procedure**

Preferably, the NMKL method should include information on the performance such as

- Selectivity (inclusivity/ exclusivity)
- Probability of detection, sensitivity, specificity, (for qualitative methods)
- Lowest validated level
- Precision (for quantitative methods)

If results from PT-schemes with the prescribed method are available, include z-score or zeta-score.

##### **3.1 Selectivity (inclusivity and exclusivity)**

For both qualitative and quantitative methods it is important to check the selectivity, i.e. if the method is able to detect the target microorganism from a wide range of strains (inclusivity) and if there are no interferences with other non-target microorganisms (exclusivity).

The extent of the testing depends on the target microorganisms, the number of available relevant strains at the laboratory, and last but not least how much effort the laboratory would like to put into the testing.

For the inclusivity, select pure cultures of target microorganisms. These strains shall be representative of the most relevant strains for the matrix of interest. From each test strain

culture an appropriate growth medium is cultured overnight. Analyse and see if the method detects the target microorganisms.

For the exclusivity, select pure cultures of non-target microorganisms chosen from both the strains known to cause interference with the target microorganism and from strains naturally present in the food matrices. From each test strain an appropriate growth medium is cultured overnight before the analysis.

### 3.2 Samples in the internal study

It is important to know the approximate level of contamination, and hence artificial contamination of samples might be used. If validation is requested for all food matrices, select 3 relevant food matrices. For each matrix, use four levels:

Level 1: 0= negative control,

Level 2: 1-10 cells per x\* g (ml) sample,

Level 3: 10-100 cells per x\* g (ml) sample

Level 4: 100-1000 cells per x\* g (ml) sample

\* In qualitative analyses: x is often 25, in quantitative analyses: x is often 10

When possible, use two strains relevant for the matrix.

It might be necessary to stress the microorganisms by heating, freezing or chilling of the samples at selected time/temperature combinations. Preparation of samples and stressing strains are described in the draft ISO 16140-2 annex C.

### 3.3 Qualitative methods: Estimation of the probability of detection (POD), sensitivity, specificity and detection level

Sensitivity and specificity can be established as described in ISO 16140 and in NordVal protocol, by using "expected value" instead of "reference method", or it can be estimated according to the simplified description given below. In this protocol, the probability of detection (POD) is included. POD is used in the review of data from qualitative microbiological and chemical qualitative analyses in AOAC International and is included as optional in the new revision of ISO 16140.

Analyse 4 replicates for each level and matrices (4 replicates x 4 levels x 3 matrices = 48 analyses).

**For each level and matrix**, calculate the number of positives divided of the total number of samples (here: 4), i.e. probability of detection (POD), for the method at the specified level. In the table below, an example of results are given

*Table 1: Example of results, POD, obtained at different levels*

Level	1	2	3	4
<b>POD meat</b>	0/5 = 0	1/5 = 0,2	2/5 = 0,4	5/5 = 1
<b>POD dairy</b>	1/5 = 0,2	2/5 = 0,4	1/5 = 0,2	3/5 = 0,6
<b>POD fish</b>	0/5 = 0	2/5 = 0,4	2/5 = 0,4	4/5 = 0,8
<b>POD vegetables</b>	0/5 = 0	1/5 = 0,2	2/5 = 0,4	4/5 = 0,8
<b>POD pastry</b>	0/5 = 0	1/5 = 0,2	2/5 = 0,4	3/5 = 0,6

Specificity = 1 - POD, for the results of level 1 (blind control).

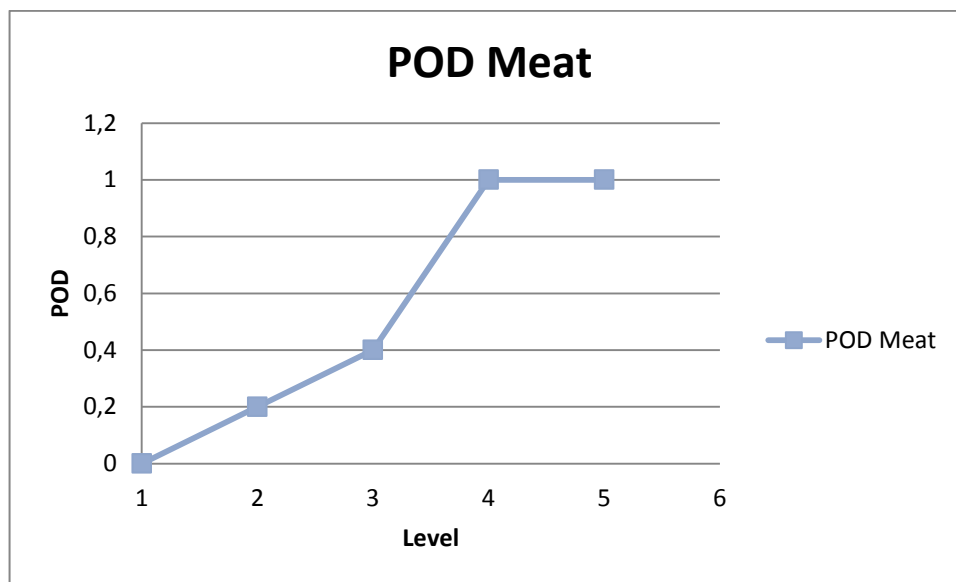
Specificity is the number of negatives that are expected to be negative

Sensitivity = POD, for the results of level 2-5

Sensitivity is the number or positives that are expected to be positives. For levels close to the detection level, the sensitivity is normally about 0,5 (50%). In the example given, level 3 would be below the detection level, while level 4 would be above the detection level. Here, the detection level would be below 1000 cfu/25g. The sensitivity depends on the level. Method A with a sensitivity of 100% is not necessarily better than a Method B with a sensitivity of 50%, i.e. if the sensitivities are estimated at different levels. For example if Method B achieves a sensitivity of 50% at level 2, but would reach 100% sensitivity at level 3, and the 100% sensitivity for Method A derived from level 4, but would have given 30% at level 3. Then it is actually Method B that is more sensitive than Method A.

For each matrix, illustrate it graphically, as e.g. the POD for meat in table 1.

Figure 1: POD-curve for meat, from table 1



The POD curve is the curve of the sensitivity.  
The specificity is  $1 - \text{POD}(0)$ , in this case 1 (100%).

#### *Drawing conclusions*

Based on the POD-curve for meat products, the method tested is suitable for levels 100-1000 cfu/25 g (level 4) and above.

Detection level are often defined as the level where the sensitivity is 50%, i.e.  $\text{POD} = 0,5$ , which for meat would be between level 3 and 4.

False positives, the probability of obtaining positive response when it should have been negative, is where the POD curve cuts the y-axis,  $\text{POD}(0)$ .

False negatives, the probability of obtaining negative response when it should have been positives, is the difference between  $\text{POD} = 1$  and the POD curve,  $1 - \text{POD}(\text{level})$ .

### 3.4 Quantitative methods: Reliability of the method

For each level and food category (see 3.2) analyse in duplicates or use different dilutions (however in well countable range, not too low or too high). Repeat the analyses on another day/ or by another analyst in order to get internal reproducible conditions. For level 2-5, calculate the mean, M, (in log cfu/g) and the standard deviation, SD (in log cfu/g) for the duplicate performed on the different day and calculate the median of these again as illustrated in the table below.

Table 2: Calculation of the mean and the precision

Day / Analyst	Replicate a	Replicate b	w = a-b	w <sup>2</sup>	y=(a+b)/2
1	a <sub>1</sub>	b <sub>1</sub>	w <sub>1</sub> =a <sub>1</sub> -b <sub>1</sub>	w <sub>1</sub> <sup>2</sup> = w <sub>1</sub> · w <sub>1</sub>	y <sub>1</sub> =(a <sub>1</sub> +b <sub>1</sub> )/2
2	a <sub>2</sub>	b <sub>2</sub>	w <sub>2</sub> =a <sub>2</sub> -b <sub>2</sub>	w <sub>2</sub> <sup>2</sup> = w <sub>2</sub> · w <sub>2</sub>	y <sub>2</sub> =(a <sub>2</sub> +b <sub>2</sub> )/2
Calculate the sum of the w <sup>2</sup>				w <sup>2</sup> =w <sub>1</sub> <sup>2</sup> +w <sub>2</sub> <sup>2</sup>	
Calculate the repeatability, s <sub>r</sub> <sup>2</sup>				w <sup>2</sup> /2	
Calculate the mean (here also the median)					m = (y <sub>1</sub> +y <sub>2</sub> )/2
Calculate the standard deviation of y				s <sub>y</sub>	
Square the standard deviation of y				s <sub>y</sub> <sup>2</sup>	
Calculate the between laboratory variance, s <sub>L</sub>				s <sub>L</sub> <sup>2</sup> = s <sub>y</sub> <sup>2</sup> - (s <sub>r</sub> <sup>2</sup> /2)	
Calculate the variance of the reproducibility				s <sub>R</sub> <sup>2</sup> = s <sub>L</sub> <sup>2</sup> + s <sub>r</sub> <sup>2</sup>	
Calculate the standard deviation of the repeatability, s <sub>r</sub> , and the standard deviation of the reproducibility, s <sub>R</sub>				s <sub>r</sub> = √s <sub>r</sub> <sup>2</sup> s <sub>R</sub> = √s <sub>R</sub> <sup>2</sup>	
Calculate the limit of the repeatability, r, and the limit of the reproducibility, R				r = 2.8 · s <sub>r</sub> R = 2.8 · s <sub>R</sub>	

#### *Drawing conclusions*

As a rule of thumb the standard deviations should not be more than 0.4 log cfu/g. Often, the between laboratory variance, S<sub>L</sub> is negligible, and hence the standard deviation of repeatability and standard deviation of reproducibility are about the same. State the lowest validated level with satisfactory precision, i.e. the level where the s<sub>R</sub> is no more than 0.4 log cfu/g.

### 3.5 Use of results from PT-schemes

Traditionally, microbiological laboratories calculate z-scores when following up on their PT-results. The z-score is calculated as follows:

$$z - score = \frac{x_{lab} - x_m}{s}$$

Where

$X_{lab}$  is the obtained value at the laboratory

$X_m$  is the reference value (the mean of the participants results), and

$s$  is the standard deviation of the reference value

In z-score, the standard deviation of own results are not included.

Thus, zeta-score, which includes the uncertainty obtained at own laboratory with the method to be tested, should preferably be used.

$$zeta - score = \frac{x_{lab} - x_{ref}}{\sqrt{u_{lab}^2 - u_{ref}^2}}$$

Where

$X_{lab}$  is the obtained value at the laboratory

$X_{ref}$  is the reference value (the mean of the participants results), and

$u_{ref}$  is the uncertainty of the reference value (standard deviation of the participants results divided by the square root of the number of participants can also be used)

$u_{lab}$  is the uncertainty obtained by the laboratory it selves.

(How to estimate the uncertainty is described in e.g. NMKL Procedure No. 8)

The zeta score for the method should be no more than  $\pm 2$ .

### 3.6 Reporting of the results in the method

Report the results in an Annex to the method. Describe which matrices, strains and levels that have been tested and state the results in tables.

## 4. **References**

AOAC – 1999: Methods Committee Guidelines for Validation of Qualitative and Quantitative Microbiological Methods.

ISO/DIS 16140 – 2003: Protocol for the Validation of Alternative Microbiological Methods.

NordVal Protocol 2009: Protocol for the validation of alternative microbiological methods

NordVal Protocol No. 2, 2010: Guide in validation of alternative proprietary chemical methods

NMKL Procedure No 8, Version 4, 2008: Measurement of uncertainty in quantitative microbiological examination of foods.

NMKL Procedure No 20, 2007: Evaluation of results from qualitative method.