

## **Kolloquium „Statistische Methoden in der empirischen Forschung“**

Wann: 26. Januar 2016, 17:00 – 18:30 Uhr

Wo: Robert Koch-Institut | Nordufer 20 | 13353 Berlin (Wedding),  
S41, S42, U9 Westhafen | U9, Bus 142 Amrumer Str

**Sofia Duarte (National Food Institute, DTU, Lyngby)**

***In collaboration with Anders Stockmarr and Maarten Nauta***

### **A new method to fit a distribution to microbial counts: making sense of zeroes**

The accurate estimation of prevalence and concentration of microorganisms in foods is an important element of quantitative microbiological risk assessment (QMRA). The uncertainty in the analysis of microbial data can be reduced by performing a single-step characterization of contamination. This is possible by using a method that estimates both prevalence and concentration from the same set of quantitative enumeration data, hence avoiding the need for collection of detection data and its combined analysis with enumeration data.

Microbial concentrations are usually described by a probability distribution representing variability, based on enumeration data. Among these data often occur artificial zero counts, which originate by chance from contaminated sample units. When these are not differentiated from true zero counts from uncontaminated units, estimates of prevalence and concentration may be inaccurate. This inaccuracy is especially relevant in situations where highly pathogenic bacteria are involved and where growth can occur along the risk pathway.

A method was developed in R software that estimates the variability distribution of concentrations and differentiates between artificial and true zeroes, thus also allowing for the estimation of prevalence. It was developed and evaluated with in silico generated data following the assumptions that plate counts are Poisson distributed, that the microorganisms are homogeneously distributed within each sample unit, that there are no measurement errors during enumeration and that the choice of the parametric distribution to describe the variability of concentrations (Poisson-lognormal) is correct. We consider a lot of food units with true prevalence of contamination  $p$ . The  $\log_{10}$  of the concentration of microorganisms in contaminated units follows a normal distribution, with mean  $\mu$  and standard deviation  $\sigma$ . The method uses raw plate count data as input and does not assume a limit of quantification (LOQ) for interpretation of zero counts. Furthermore, counts resulting both from contaminated and uncontaminated units are analysed together and the proportion of artificial zeroes is estimated.

The overall performance of the method is satisfactory for the interpretation of microbial counts under different conditions of true prevalence and concentration, with the best estimates obtained for  $p$  and negligible bias observed in the estimation of  $\mu$  and  $\sigma$ . It performs particularly well at low  $p$  levels and low expected  $\sigma$ .

This new method represents an analytical tool to be used for the analysis of microbial enumeration data with the possibility to be adapted for the future interpretation of

quantitative results obtained with alternative enumeration techniques that start to be used in food microbiology, such as enrichment real-time PCR. It has been published

During the presentation, the theoretical basis of the method, as described in Duarte et al., 2015, "Fitting a distribution to microbial counts: Making sense of zeroes" (International Journal of Food Microbiology, Vol. 196), will be explained. Furthermore, a demonstration using in silico data will be performed, following the script available in appendix A of the above mentioned publication.