

Handling of left- and interval-censored Arsenic data from BfR MEAL study for dietary Exposure Assessment

Annett Martin, Irmela Sarvan, Matthias Greiner
Bundesinstitut für Risikobewertung (BfR), MEAL-Studienzentrum, Max-Dohrn-Strasse 8-10, 10589 Berlin

Introduction

The BfR MEAL study is Germany's first Total Diet Study. The study aims to provide representative summary statistics for substance concentrations in 356 foods, typically consumed by the population in Germany for the assessment of dietary exposure. A difficult step in dietary exposure assessment is the handling of concentration data reported as censored. The simple substitution can lead to biased estimates. We are interested in a correction of bias due to the censoring and estimation of confidence limits for the statistics. The latter or specialised statistical approaches may be used to support group comparisons.

Sample size and methods

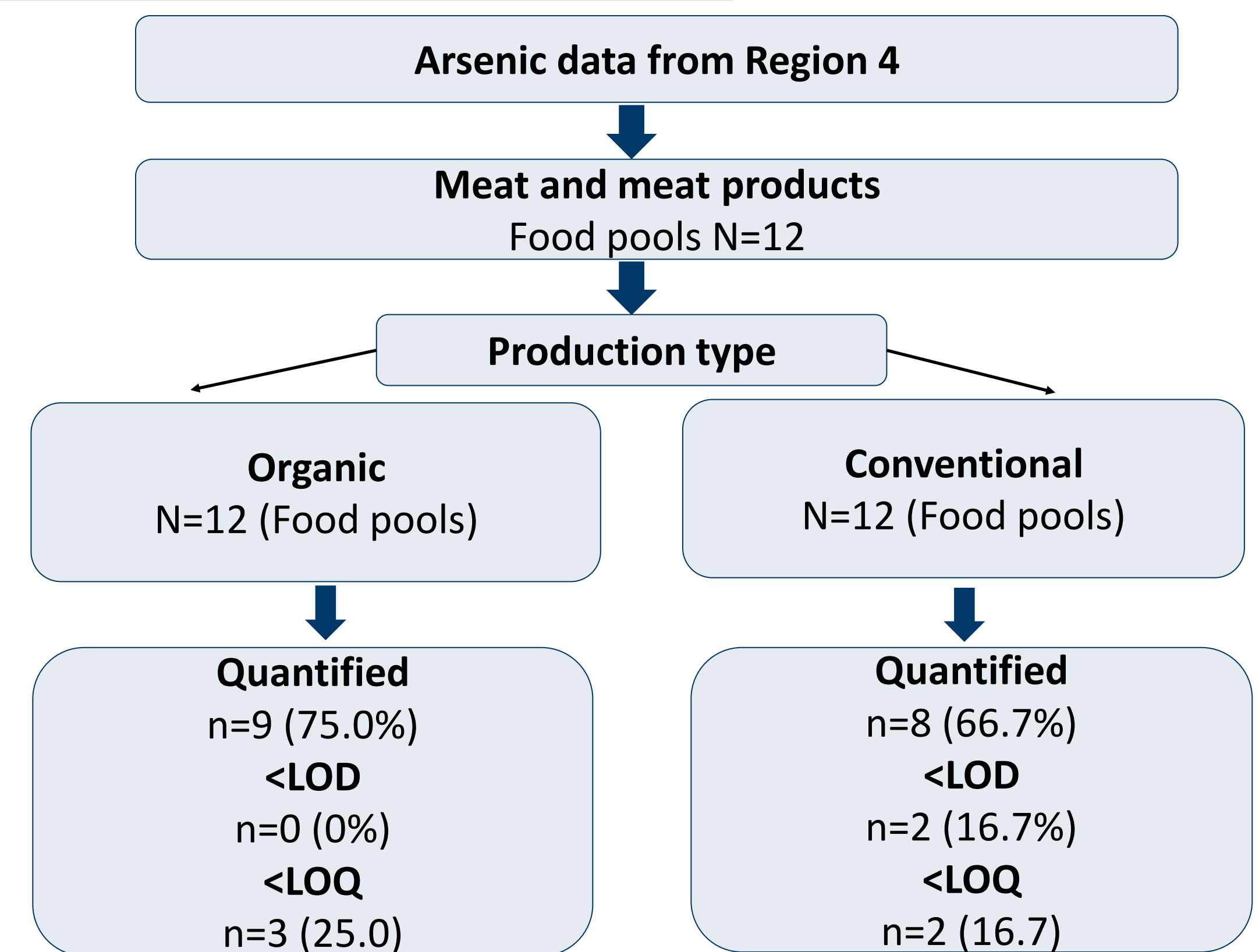
Examples for three different result values for arsenic data [mg/kg]

Main food group	Food pool	Production type	LOD	LOQ	Value type		mLB (left)	UB (right)
Meat and meat products	Pork liver, Cooked	Conventional	0.001	0.002	<LOD	left	0 (NA)	0.001
Meat and meat products	Chicken, cooked	Organic	0.001	0.002	<LOQ	Interval	0.001	0.002
Meat and meat products	Poultry liver	Conventional	0.001	0.002	Result	Quantifiable	0.004	0.004

1. *Left censoring* – a data point is below LOD, but it is unknown by how much (nondetects)
2. *Interval censoring* – a data point is on an interval between LOD and LOQ
3. *Quantifiable* - uncensored

mLB Mean = 0.0017 Median = 0.001
UB Mean = 0.0023 Median = 0.002

3 Value types



2 steps for estimation of Arsenic concentration in foods with R

1. "fitdistrplus" package

- Median estimation based on maximum likelihood
- Fitting algorithm accounted for the censored observations (lognorm)
- Nonparametric Bootstrap resampling in order to simulate 95% uncertainty in the parameters of the distribution

2. "mc2d" package

- two-dimensional Monte-Carlo (MC) simulation to simulate final parameter estimation
- Variability and uncertainty in the parameters may be estimated separately

Significance testing for group comparisons

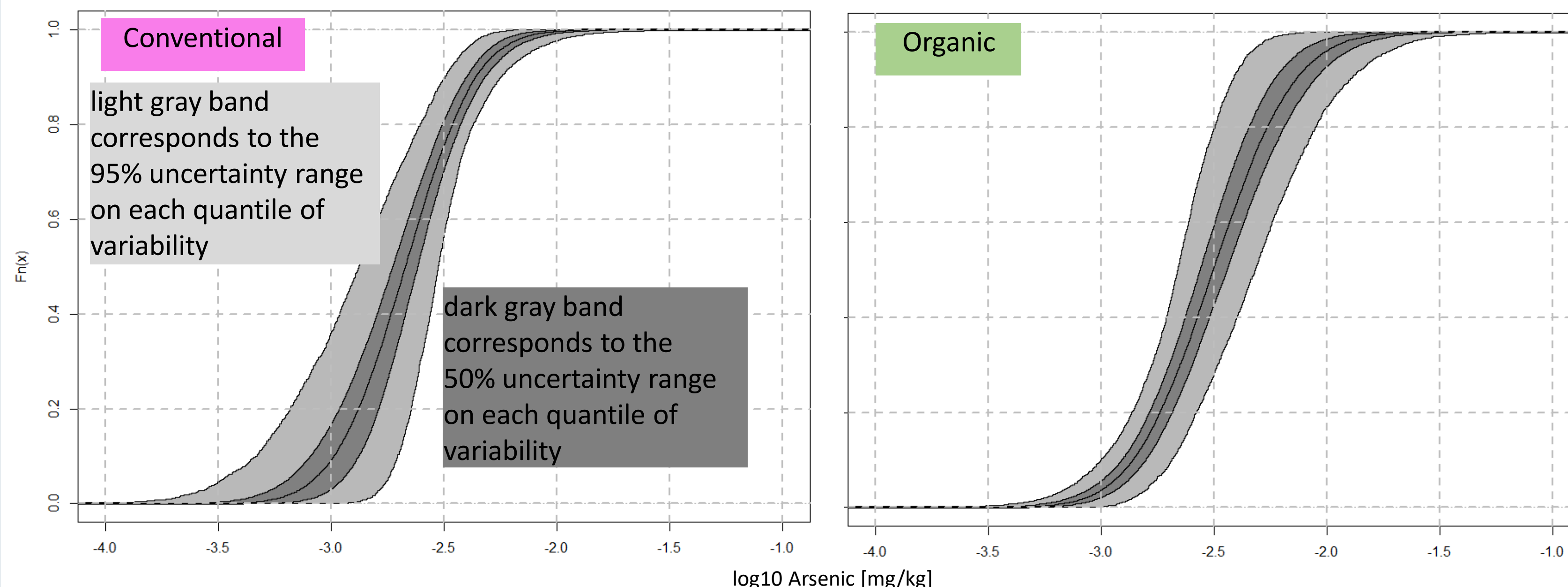
"interval" package

- Testing differences in As between organic and conventional produced foods
- non-Parametric maximum likelihood estimation (NPMLE)-Wilcoxon exact test
- 95% CI for p values (exact permutation test)

Results

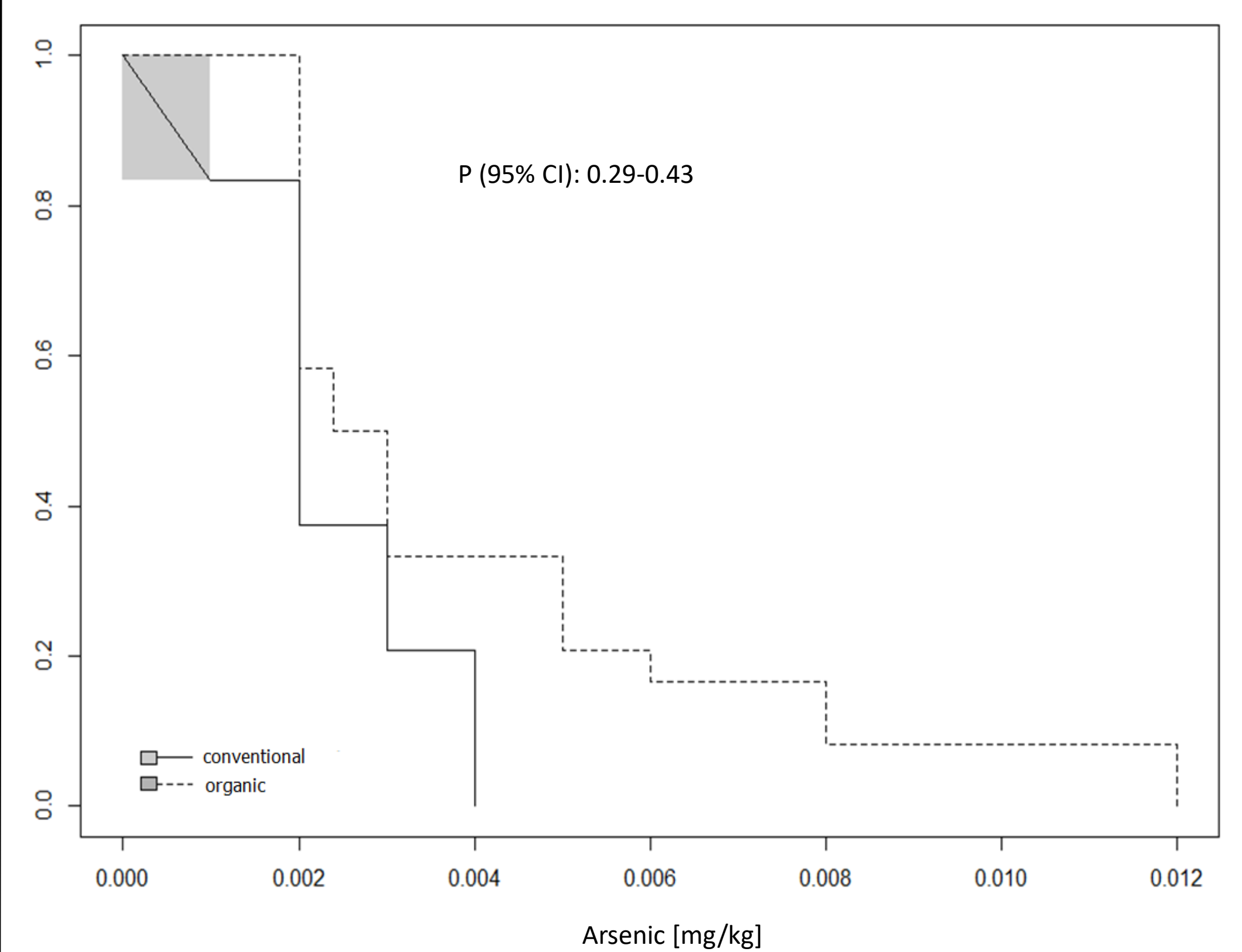
Estimation of Arsenic concentration with Lower & Upper bound and two-dimensional Monte Carlo simulation

Production type	No of food pools N (%)	mLB (mg/kg)			UB (mg/kg)			MC			
		Median	Min	Max	Median	Min	Max	Median	95% Uncertainty	Min	Max
Conventional	12 (50)	0.0025	0	0.004	0.0025	0.001	0.004	0.0021	0.0013; 0.003	0.00035	0.012
Organic	12 (50)	0.0027	0.001	0.012	0.0027	0.002	0.012	0.0032	0.0022; 0.0048	0.0004	0.025



Cumulative distribution plots of arsenic in meat and meat products and area of uncertainty

With the function ictest (from package interval) no differences between organic and conventional produced food (meat and meat products) were observed.



Estimated cumulative distribution plot of two groups. The dashed curves correspond to the organic group; the solid curves correspond to the conventional group.

References

- Fay MP, Shaw PA (2010). "Exact and Asymptotic Weighted Logrank Tests for Interval Censored Data: The interval R Package." *Journal of Statistical Software*, *36*(2), 1-34. doi:10.18637/jss.v036.i02 <<https://doi.org/10.18637/jss.v036.i02>>.
- Delignette-Muller ML, Dutang C (2015). fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software*. <https://www.jstatsoft.org/article/view/v064i04> DOI 10.18637/jss.v064.i04.
- Pouillot R, Delignette-Muller ML. Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. *Int J Food Microbiol*. 2010 Sep 1;142(3):330-40. doi: 10.1016/j.ijfoodmicro.2010.07.011. Epub 2010 Jul 15.