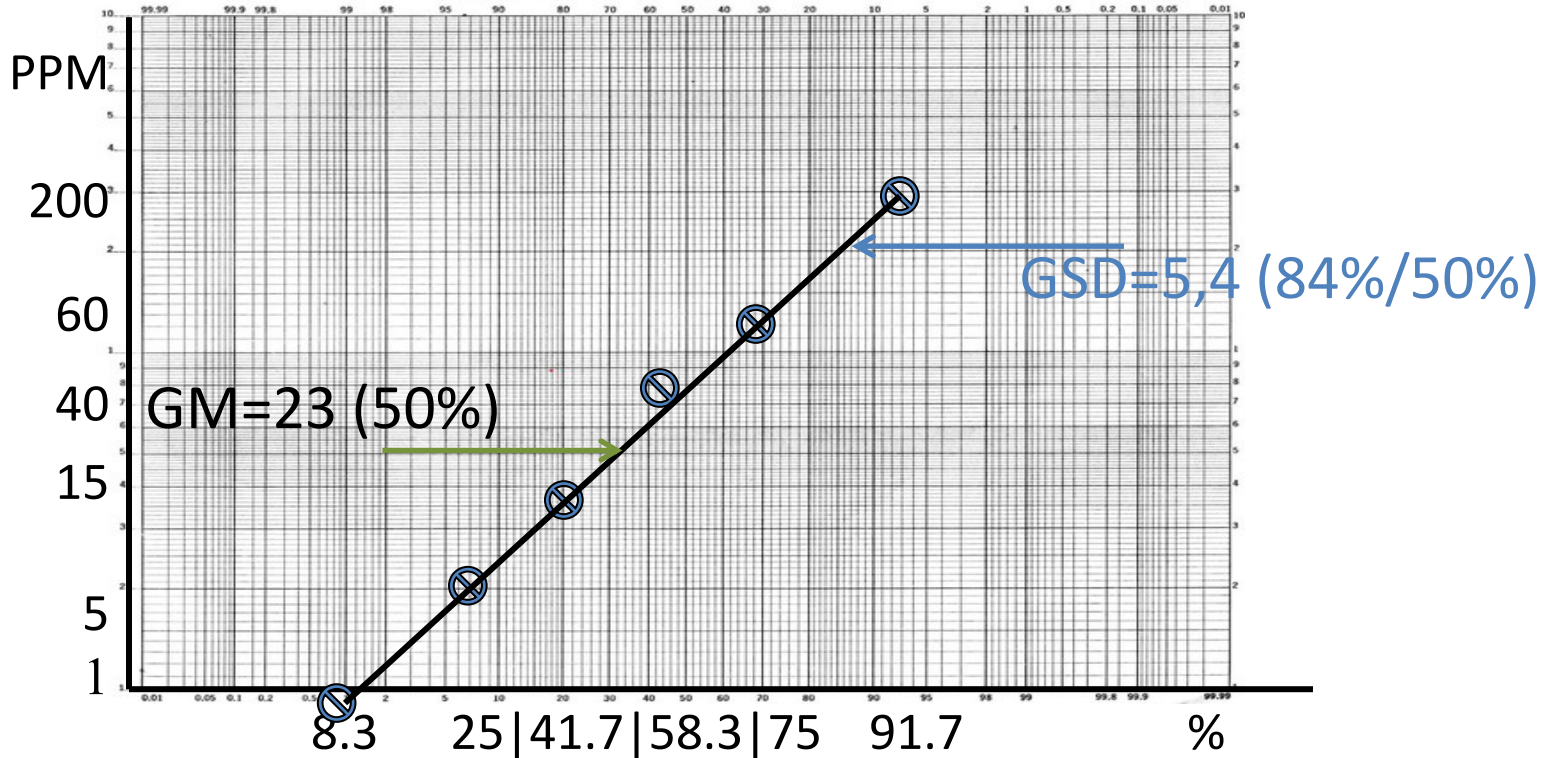# Examining exposure series

1. Lognormal goodness-of-fit
   – visual, test & transformations
2. Processing undetectables
   – (fraction LoD), regression, degrees of freedom
3. GSD values (5.4.2)
   – Too low, too high, all is possible
4. Individual outliers
   – Location, dispersion, between and within
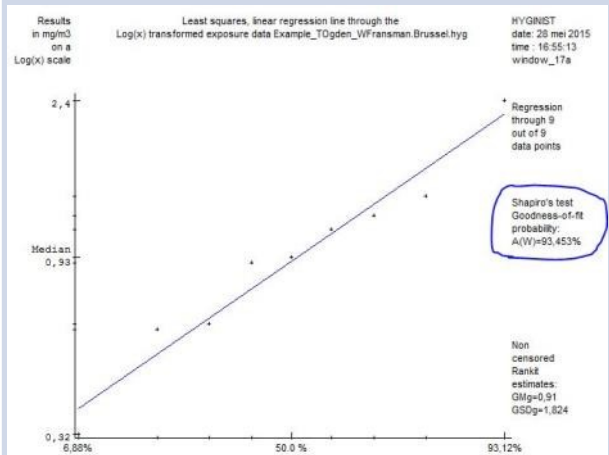
LogNormal probability paper and NOS regression

Annex E .1.1 through 1.4

Lognormal plot and Shapiro test
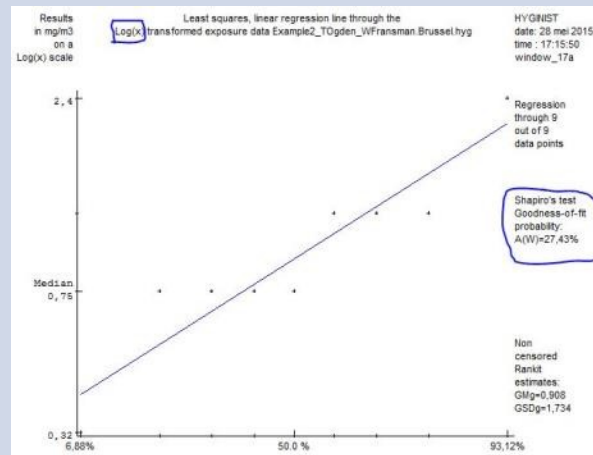
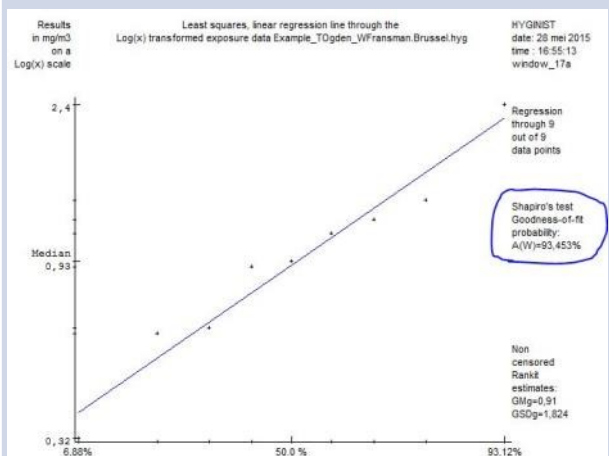| Continuous outcome	Discrete data
P_{SW}	P_{SW}
Single outlier	Detection limit
P_{SW}	P_{SW}

Annex E.1 and E.2 In series up to 9 samples, omnibus goodness-of-fit tests (including Shapiro) will hardly reject Lognormality based on a 5% criterion
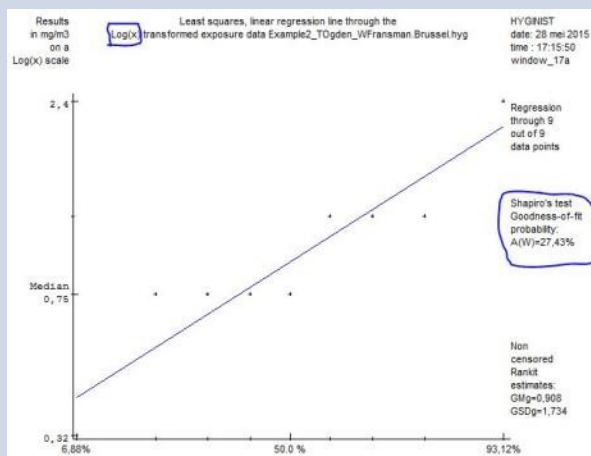
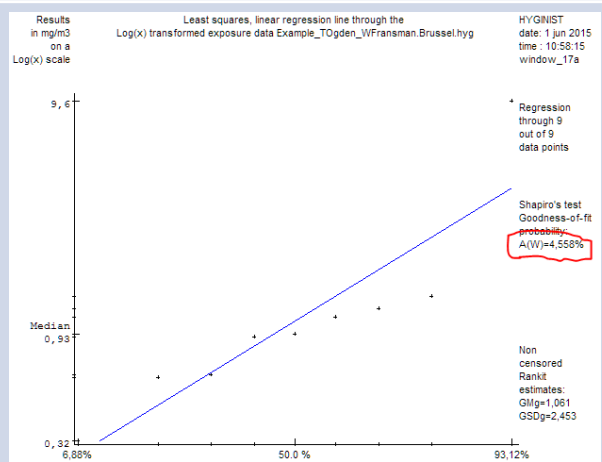# Goodness-of-fit plot and test (Annex E.1 and E.2)

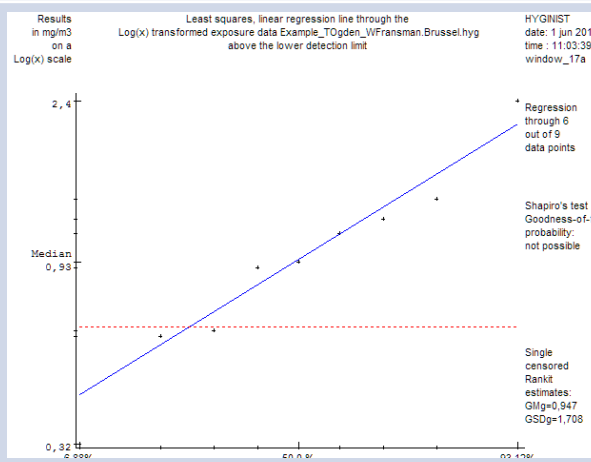| Continuous outcome	Discrete outcome

P_{SW}=93%

P_{SW}=27%

Single outlier (9,6 in stead of 2,4)

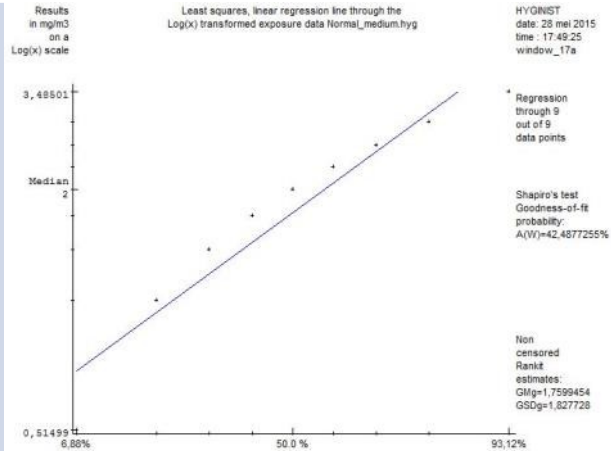### Detection limit

P_{SW}=4.5%

P_{SW}=n.a.

In series up to 9 samples, plot and omnibus tests (Shapiro by preference) help to validate SEG and Lognormality

Compare transformations
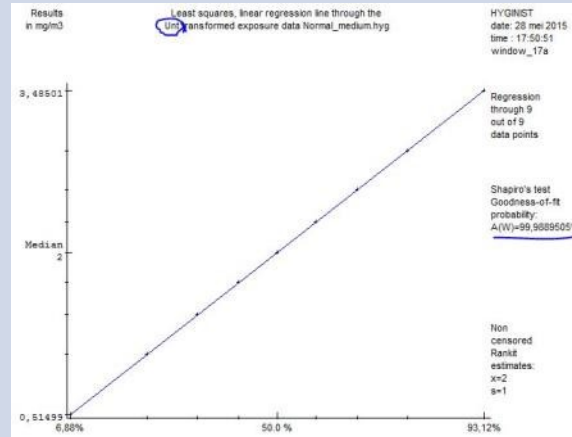
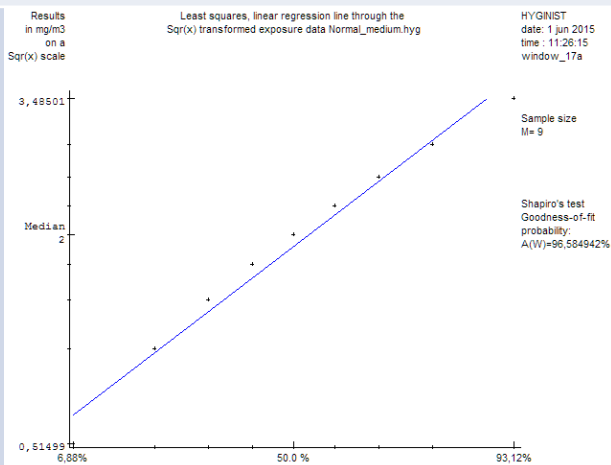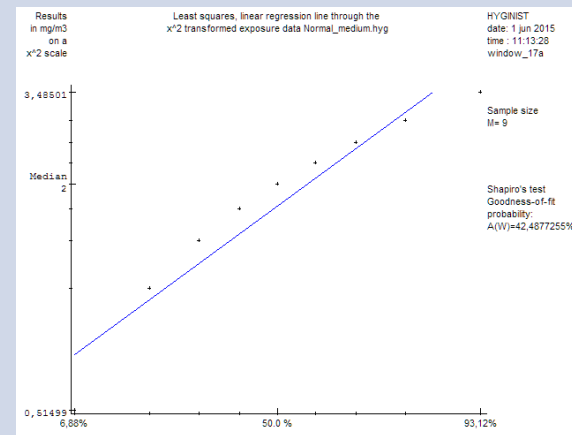| log	untransformed
$P_{SW}=43\%$	$P_{SW}>99.9\%$
Square root	Squared
$P_{SW}=97\%$	$P_{SW}=42\%$

Annex E.1 and E.2 Graphical and statistical tests help to assess if the exposure distribution can be represented by the lognormal model.

Recommendations 5.4 & Annex E.3.3

- Lognormality is the null hypothesis (H_0)
- to accept or reject H_0 use for sample size (N=3->9):
  - Probability paper, and
  - An omnibus test (preference Shapiro), but no criterion, and
  - Compare different transformations (un-, log-, etc.)
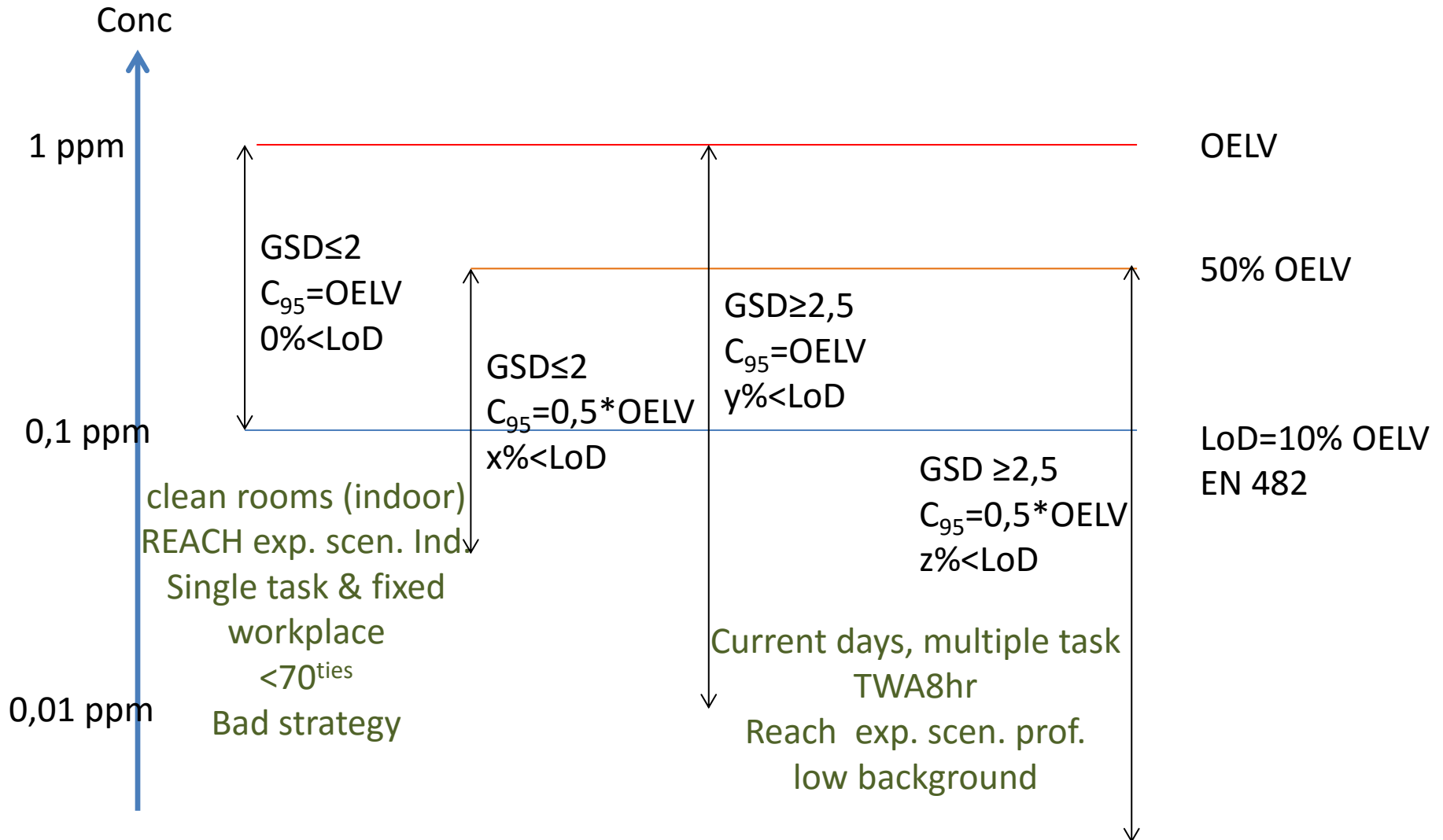- For higher sample size and nondetects consult expert

Examining exposure series

1. Lognormal goodness-of-fit
   - visual, test & transformations

2. Processing undetectables
   - (fraction LoD), regression, degrees of freedom

3. GSD values
   - Too low, too high, all is possible

4. Individual outliers
   - Location, dispersion, between and within

The urgency to treat undetectables more professionally (~~0,5*LoD~~)
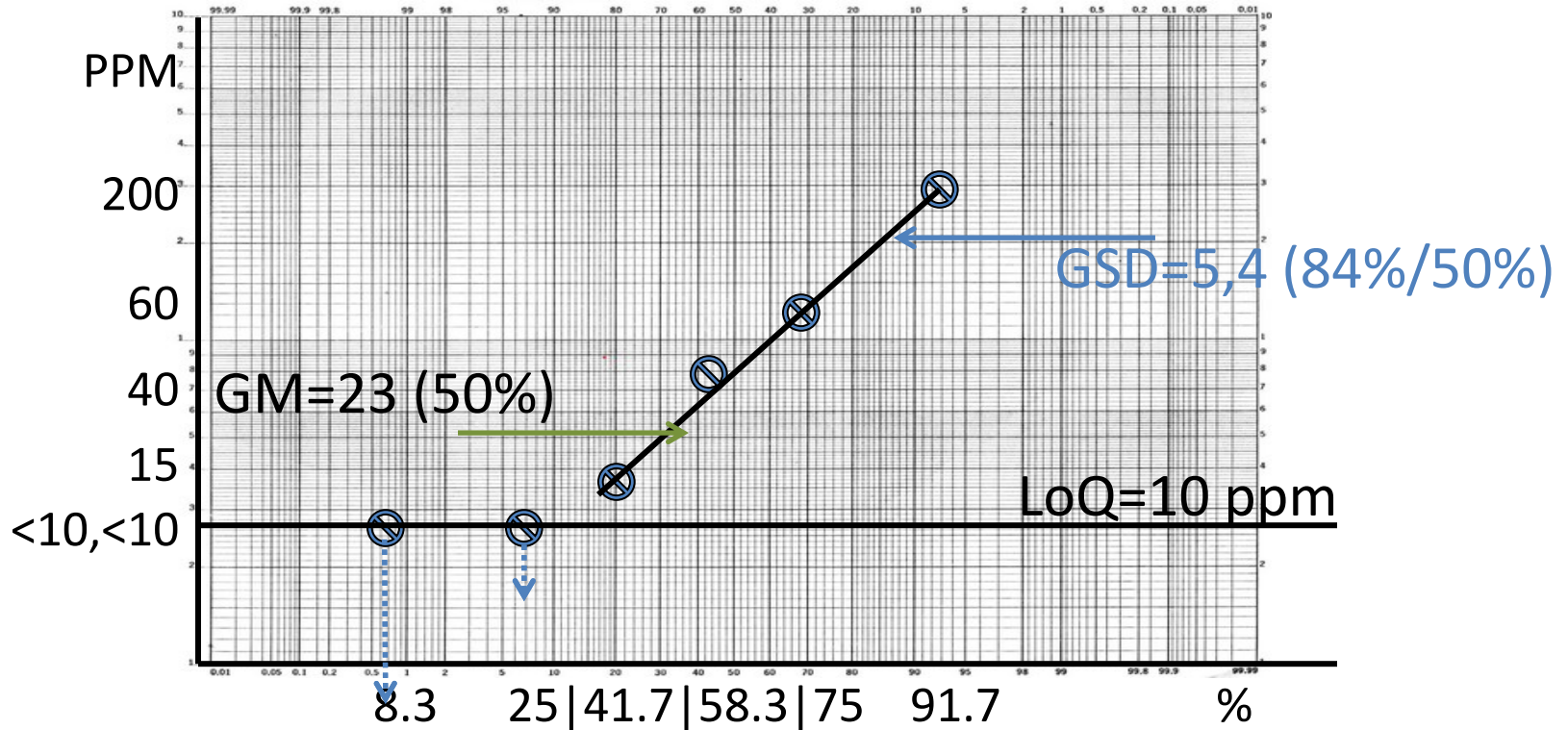
- Overestimation of GM
- Underestimation of GSD and $C_{95\%}$
- Changing working patterns
- Trends in time
- IH reputation

The urgency to treat undetectables more professionally (~~0,5*LoD~~)

Conc

1 ppm ———————————————————————— OELV

GSD≤2
C_{95}=OELV
0%<LoD

50% OELV

GSD≥2,5
C_{95}=OELV
y%<LoD

GSD≤2
C_{95}=0,5*OELV
x%<LoD

0,1 ppm

LoD=10% OELV
EN 482

GSD ≥2,5
C_{95}=0,5*OELV
z%<LoD

clean rooms (indoor)
REACH exp. scen. Ind
Single task & fixed
workplace
<70[ties]
Bad strategy

Current days, multiple task
TWA8hr
Reach exp. scen. prof.
low background

0,01 ppm

NOS regression through the detectables

Degrees of freedom

- If the extrapolation space increases, the number of degrees of freedom df decreases
- Small sample size (& small # detectables) increases exceedance probability

Degrees of freedom (df)

Small sample size (N) and undetectables (N-k):

- decrease df

- increase the extrapolation space

- increase noncompliance probability, in noncStudent, unbiased, nonPar methods

The influence of k<N detectables on df:

N=9, K=3

- df=k-1 (Krishnamoorthy) ,

df=5

- df=N-1 (NDExpo, BWStat)

df=8

- df=Int((N+k)/2)-1 (HYGINIST)

df=7

Non-detects in Industrial Hygiene

- Annals Occupational Hygiene (2009-2010, *Large sample* solutions):
  - Ogden. Editorial: Handling results below the level of detection.
  - Helsel. Incorporating Non-detects in Science.
  - Flynn. Analysis of censored exposure data by the Shapiro-Wilk W statistic.

- BOHS-NVvA guidance: "It is not recommended simply to substitute LoQ/2 or LoQ/√2 for each value<LoQ"

- Software (probability paper, regression):
  - HYGINIST (1990, small samples solution), Altrex, IHDataAnalyst, NDExpo, BW_Stat (2014)

NIOSH 1977: be ware of constant background exposure
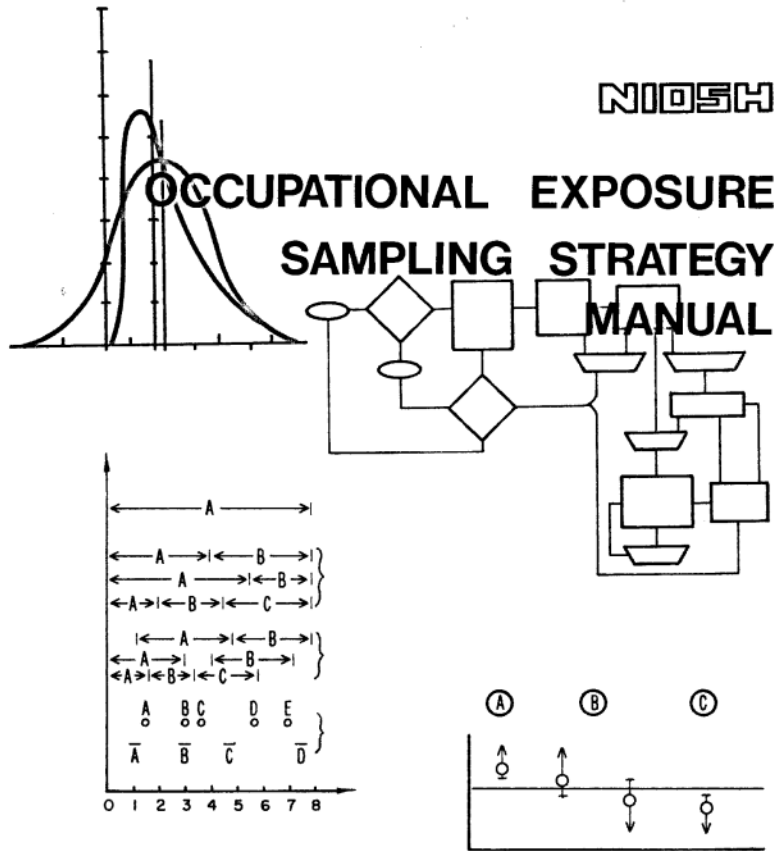
U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Center for Disease Control
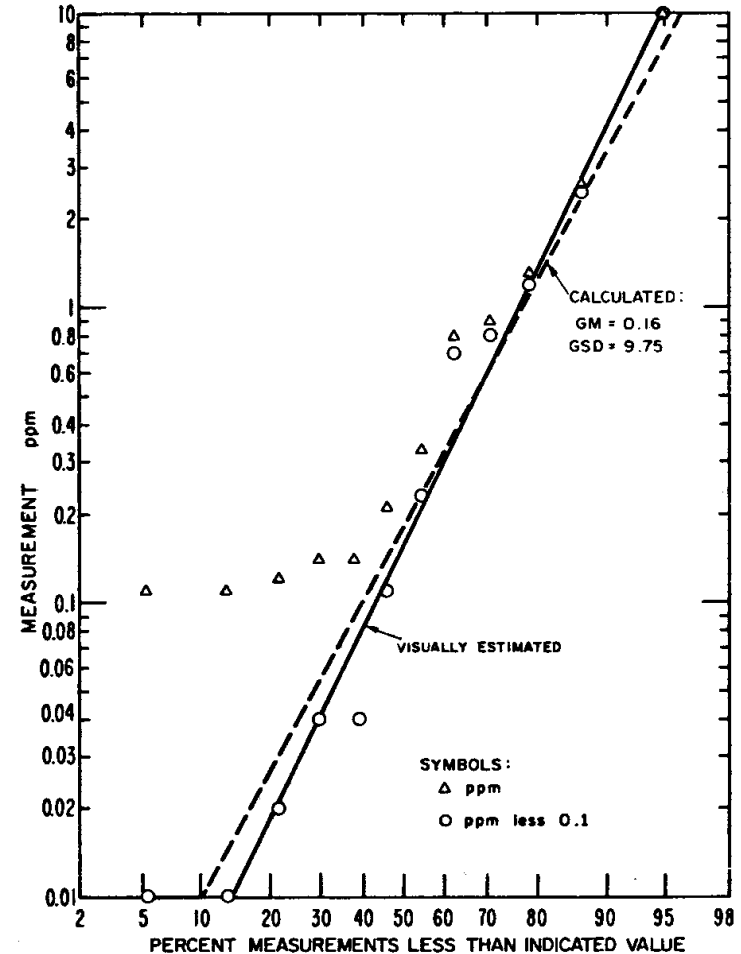National Institute for Occupational Safety and Health

Figure I-4. Hydrogen fluoride measurement distribution.

Theory censored samples

- **Helmut Schneider. Truncated and censored samples from Normal populations. Statistics: textbooks and monographs. Vol <u>70</u> (1986).**

- Galton, F. Speeds of trotting horses. Proc. Royal Soc. <u>62</u> (**1898**) 310-314.

- Fisher R.A. The truncated Normal Distribution. British Assoc. Adv. Sci. Math. Tables I, **1931** pp XXXIII

- Lotz A. Statistische Analysemethoden für linkszensierte Variablen. IPA, Bochum (2013)

- "LoQ/x" methods never recommended

Estimating GM and GSD from sampling data with undetectables

Regression through the data above LoD and optimizing GM and GSD using Shapiro & Wilks Goodness-of-Fit

HYGINIST 4.2.3

NVvA BOSH guidance
3.7 Treatment of values < LoQ

- "There are ML, regression & Shapiro & Wilks methods."
- Included in BW_Stat for group and individuals
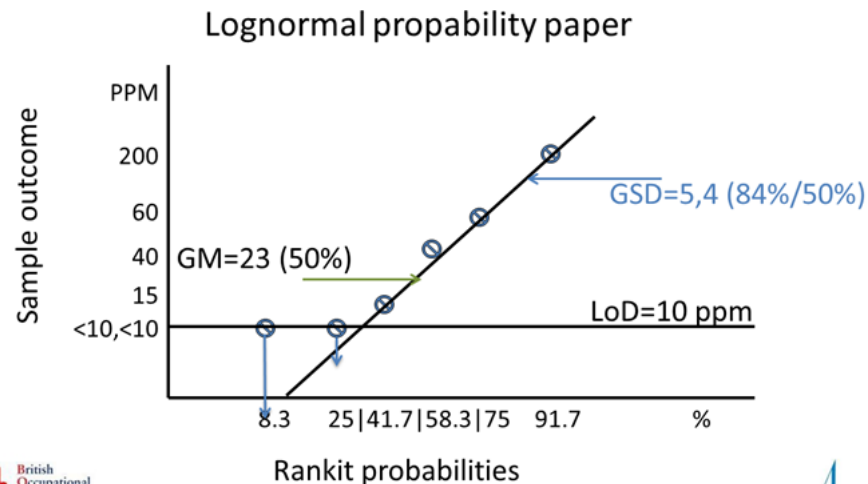

BW_Stat Startup screen

Lognormal probability plot for group and selected worker
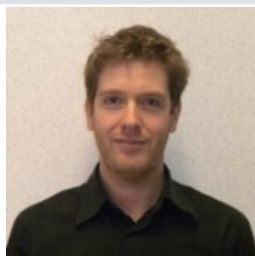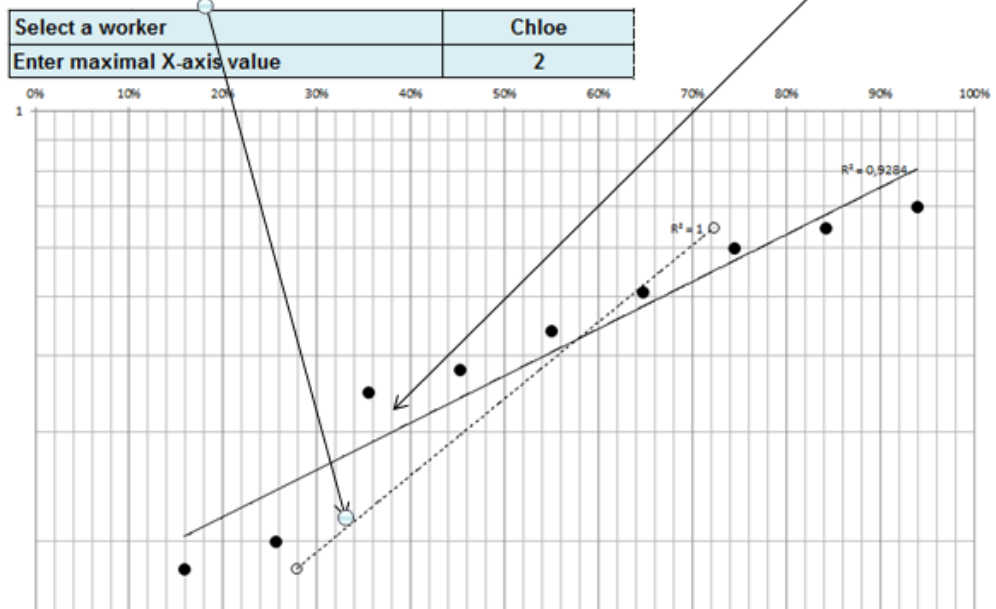
Recommendation

- a simple annex on graphical treatment of non-detects
- adjacent to/combined with Annex E proposal
- No promotion of "short cut" methods (i.e. LoQ/2)
- Refer to IH free-ware, doing the job in a reproducible way

Handle non-detects graphically

Lognormal propability paper

Proposed text

Valid measurement methods used to measure workplace air exposure must comply with CEN 482 prescribing its range includes at least 0.1 through 2 times the OELV with a x% accuracy.

This however does not prevent the occurrence of outcome outside that range. In Annex "X" some methods are given for an unbiased estimate of the exposure distribution in the presence of outcome outside the detection range

Examining exposure series

1. Lognormal goodness-of-fit
   - visual, test & transformations
2. Processing undetectables
   - (fraction LoD), regression, degrees of freedom
3. GSD values
   - Too low, too high, all is possible
4. **Individual outliers**
   - Location, dispersion, between and within