



DoE 9: Statistics and Epidemiology

ISES Europe Training Series

Ruairí Weiner, Gerald Bachler



Legal Notice

Fair Use Disclaimer

This presentation may contain copyrighted material used under the fair dealing exceptions in EU copyright law, in accordance with Directive 2001/29/EC and Directive 2019/790. It is intended solely for non-commercial, educational purposes.

Copyright Notice

© ISES Europe 2025

All rights to the original works are retained by their respective copyright holders. Any reproduction, distribution, or use of this presentation, in whole or in part, requires explicit permission from the copyright holder(s) or the author. Unauthorized use of any content is strictly prohibited.



Overview of All Training Videos

- DoE 1) Basic Concepts and Principles in Exposure Science
- DoE 2) Fundamentals of Environmental Chemistry and (Eco)Toxicology
- DoE 3) Exposure Modelling
- DoE 4) Exposure Monitoring
- DoE 5) Exposure Assessment and Risk Characterisation
- DoE 6) Risk Management and Sustainability Assessment
- DoE 7) Relevant Legislative Frameworks
- DoE 8) Risk Communication and Stakeholder Engagement
- DoE 9) Statistics and Epidemiology**

Access all videos via: <https://ises-europe.org/>



Domain of Expertise (DoE) 9: Statistics and Epidemiology

Module 1

Introduction to Statistics

Module 2

Statistics for Exposure Scientists

Module 3

Introduction to Epidemiology



DoE 9: Statistics and Epidemiology

Module 2

Statistics for Exposure Scientists

Gerald Bachler

Presenter



Meet Today's Presenter: **Gerald Bachler**

- **Current Role: Regulatory Officer | European Chemicals Agency (2024 – Present)**
- **Previous Roles:**
 - **DuPont (2019–2024):** Manager, Product Stewardship & Regulatory Affairs; Led global regulatory submissions, risk assessments, and safety initiatives.
 - **Vitis Regulatory (2017–2019):** Principal Exposure Scientist; Specialising in exposure and risk assessment for industrial chemicals.
 - **Shell (2015–2017):** Exposure Scientist/Industrial Hygienist; Handling risk assessments and ensuring REACH and CLP regulatory compliance.
- **Education:**
 - **PhD in Nanotoxicology**, ETH Zürich; **MSc in Health Care Engineering**, Technical University of Graz; **MSc in Health and the Environment**, Cranfield University.
- **Certification:** Swiss Certified Safety Engineer.



Context and Disclaimers

About This Lecture

This lecture provides an **introductory framework**, with some topics simplified for ease of understanding.

Disclaimer

- The content presented herein does not necessarily reflect the opinions, views, or positions of the presenters' employer or any affiliated organizations.
- References to specific organizations, tools, or entities are for illustrative purposes only and do not imply endorsement or critique.
- While every effort has been made to ensure the accuracy of the information presented, errors or omissions may occur.



Learning Objectives

- Understand why selecting an appropriate exposure value is crucial
- Recognize exposure distributions and the role of SEGs
- Interpret confidence intervals for exposure values
- Apply methods to handle left-censored (non-detect) data
- Identify software tools for exposure data analysis



Content

- 1. Introduction**
2. Understanding Percentiles & Distributions
3. Confidence Intervals & Data Requirements
4. Handling Left-Censored (Non-Detect) Data
5. Software Tools for Exposure Data Analysis
6. Summary

Refresher: How to Calculate Risk

Risk Characterisation Ratio (RCR):

Formula:
$$RCR = \frac{\text{Exposure Estimate}}{\text{Threshold Limit Value (e.g., DNEL, PNEC)}}$$

Margin of Safety (MoS):

Formula:
$$MoS = \frac{\text{No-Observed-Adverse-Effect Level (NOAEL)}}{\text{Exposure Estimate}}$$

→ Typically, risk assessment requires a **single exposure value** to compare against a toxicological threshold

Exposure Estimates in Practice

- In risk assessment, **not every individual can be assessed separately**
- Higher-tier tools and exposure measurements provide **distributions and/or large datasets**, not just single worst-case estimates
- **Exposure distributions** are used to represent the entire population and capture variability

ART — Exposure Predictions

Please set the options for the results.

Type of exposure

Percentile

Confidence interval

Estimate	0.71 mg/m ³
Confidence Interval	0.12 mg/m ³ –6 mg/m ³

Higher tier tool: Advanced REACH Tool

SEQN	DR11FDCD	DR11GRMS	DR11KCAL	DR11PROT	DR11CARB	DR11SUGR
130378	94000100	120	0	0	0	0
130378	94000100	120	0	0	0	0
130378	92101000	300	3	0.36	0	0
130378	94000100	240	0	0	0	0
130378	83102000	4.9	27	0.11	0.16	0.14
130378	24100020	170.1	279	45.57	0	0
130378	75113000	70	14	0.65	2.58	0.77
130378	51185000	2	9	0.22	1.27	0.09
130378	14108010	2.08	9	0.62	0.17	0.01
130378	75111000	80	13	0.5	2.36	1.34
130378	94000100	240	0	0	0	0
130378	92101000	300	3	0.36	0	0
130378	93106600	1080	263	0.77	5.9	0
130378	58160510	240	308	6.45	55.79	2.6
130378	58106330	208	510	20.26	63.29	7.09
130378	13110102	125	302	4.59	38.14	31.67
130378	94000100	360	0	0	0	0
130379	92111010	615	0	0.62	0	0

Excerpt from the NHANES dietary intake database



Content

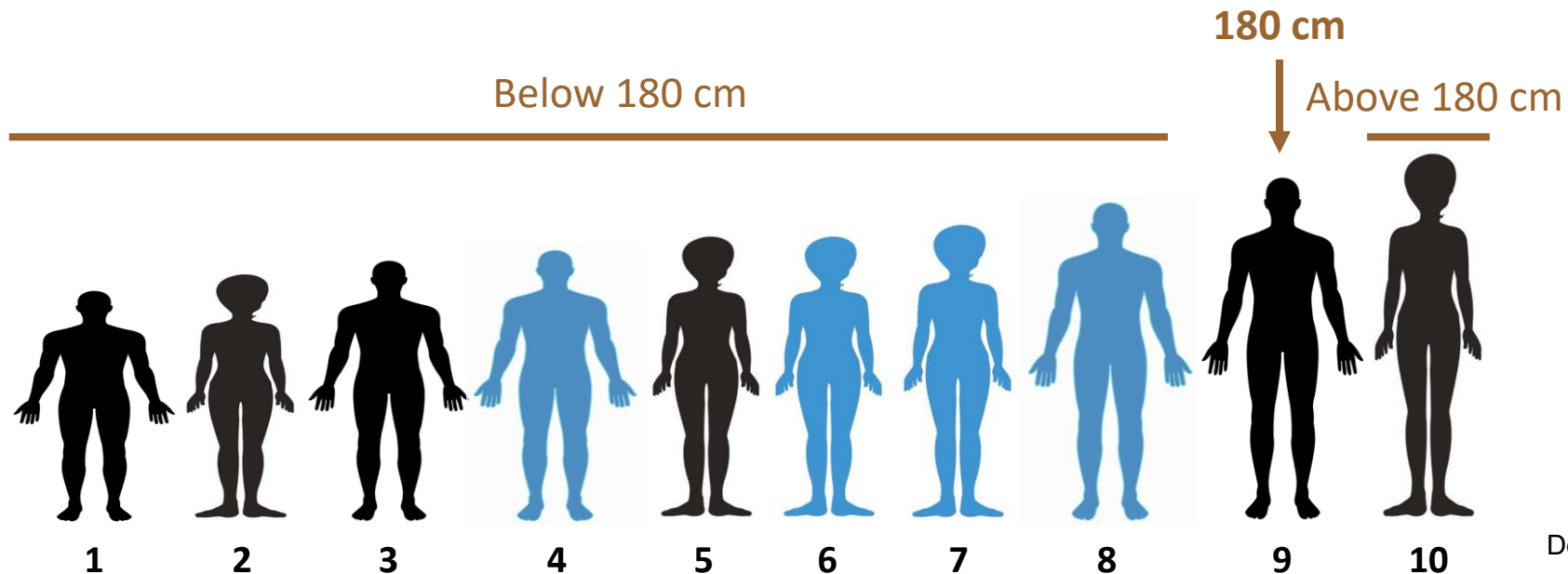
1. Introduction
- 2. Understanding Percentiles & Distributions**
3. Confidence Intervals & Data Requirements
4. Handling Left-Censored (Non-Detect) Data
5. Software Tools for Exposure Data Analysis
6. Summary

What is a Percentile

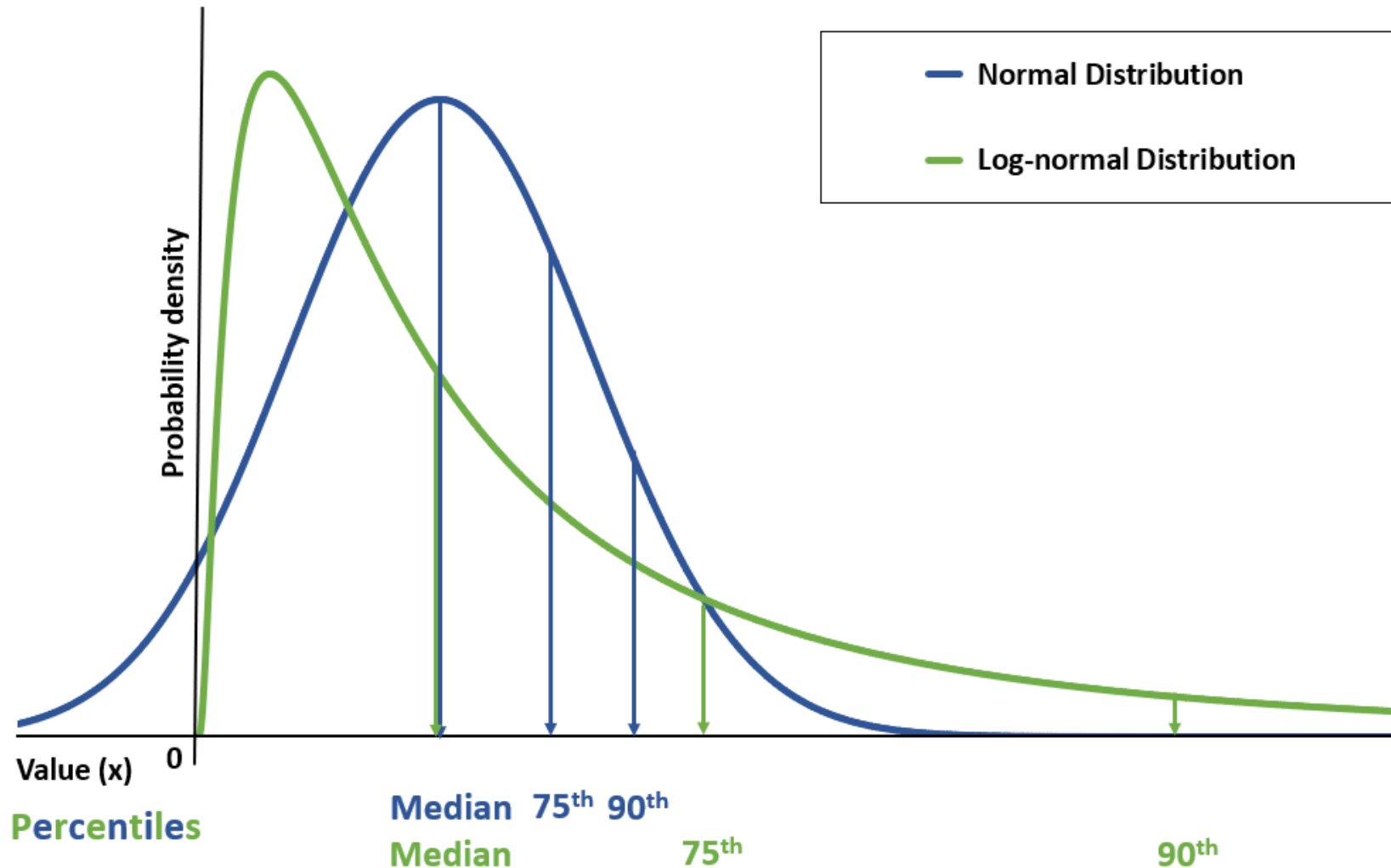
Definition: A percentile is a number denoting the position of a data point within a numeric dataset by indicating the percentage of the dataset with a lesser value.

Encyclopædia Britannica

Example: The 90th percentile height in this population is 180 cm.

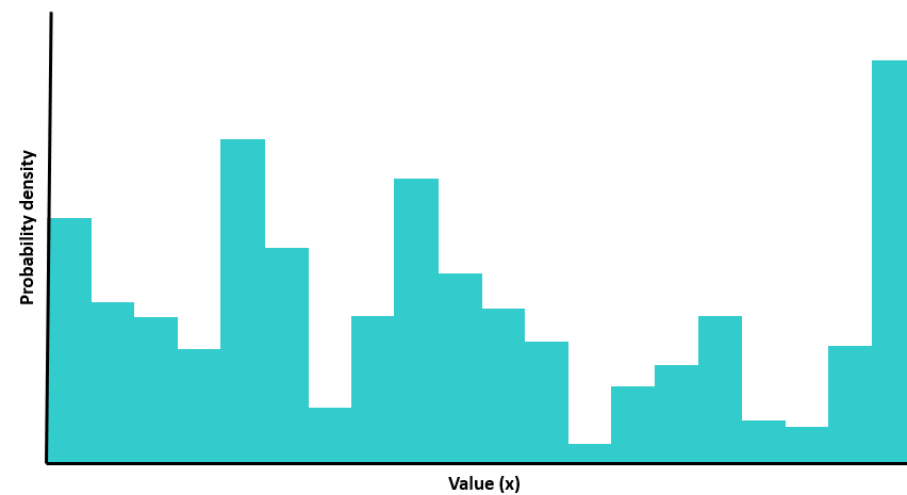
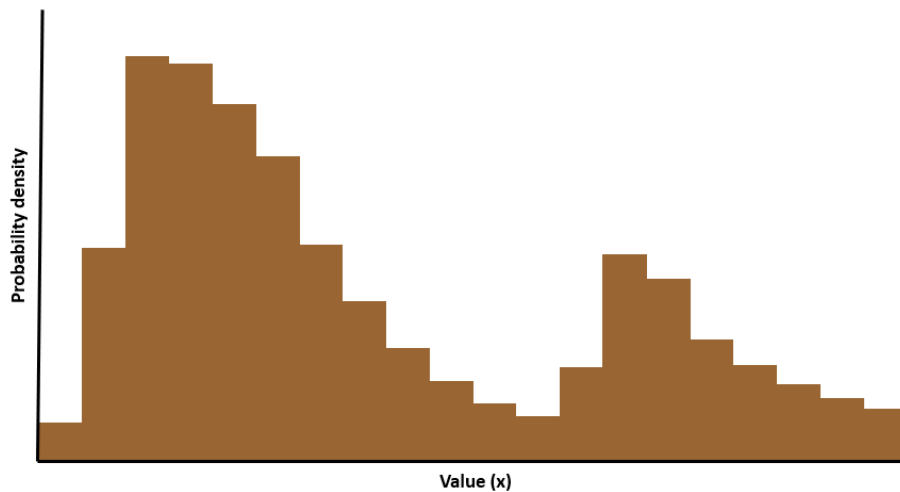
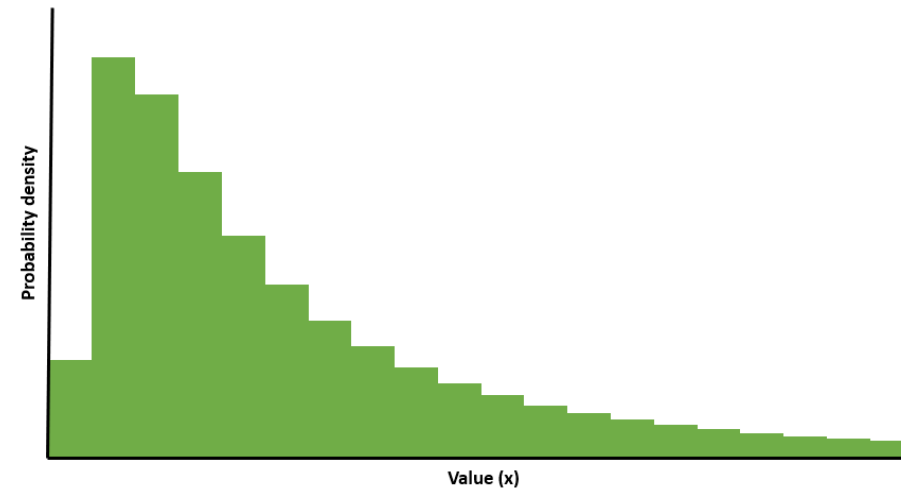
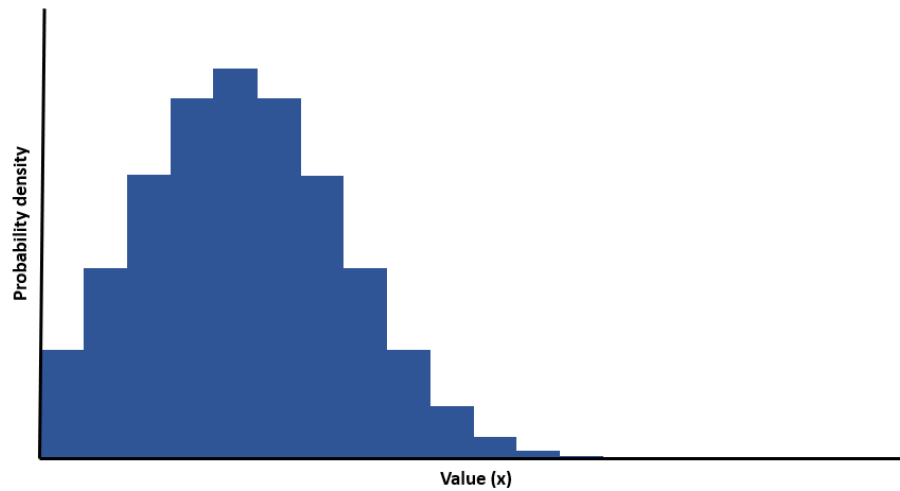


Percentiles and Distributions I



- Principle applies to normal and log-normal distributions
- **Example:** 75th percentile = 75% of measurements are below this value

Percentiles and Distributions II





Similar Exposure Group (SEG)

Definition: A **Similar Exposure Group (SEG)** is a group of receptors have the same general exposure profile for the agent being studied

Why SEGs are used:

- Group receptors with similar tasks, materials, processes, and ways of working/using products
- Enables representative sampling to estimate exposure for the whole group

Benefits:

- Avoids assessing every individual
- Identifies the nature and extent of risks for the group

How to establish SEGs:

- **Observation:** Group by task, process, and exposure patterns
- **Sampling:** Use data (historic or new) to statistically confirm similarity
- Typically assume a **log-normal distribution** for exposure



Selecting an Appropriate Exposure Value

Expert judgment is crucial: Choice of percentile strongly impacts risk assessment

Population type matters: workers, consumers, environmental receptors, vulnerable groups

Distribution type:

- Random/unknown → use more conservative percentiles
- Log-normal/established SEGs → percentile selection guided by data

Toxicological endpoint:

- Chronic → occasional high exposures may be less critical
- Acute → single high exposures may be critical

Source and quality of data:

- Peak vs. typical measurements
- High variability → may justify higher percentiles

Regulatory guidance: e.g., REACH occupational exposure guidance



Content

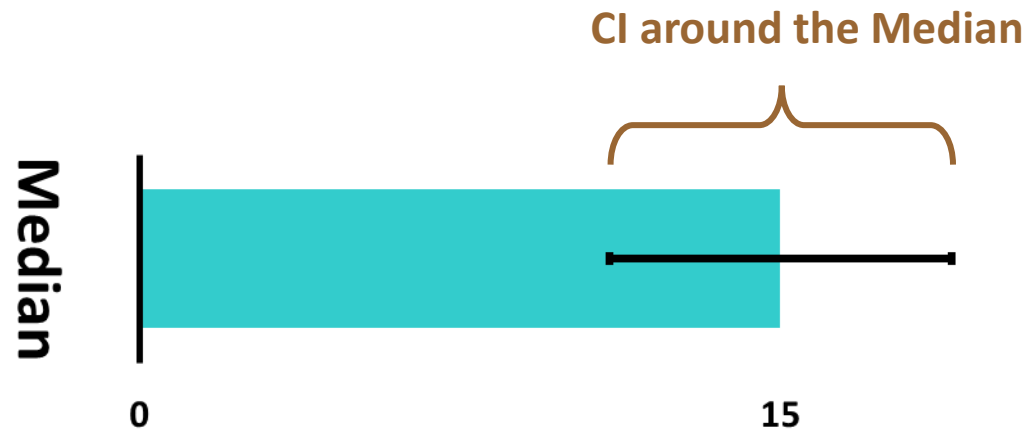
1. Introduction
2. Understanding Percentiles & Distributions
- 3. Confidence Intervals & Data Requirements**
4. Handling Left-Censored (Non-Detect) Data
5. Software Tools for Exposure Data Analysis
6. Summary

Confidence Interval (CI)

Definition: *The confidence interval (CI) shows the range of values you expect the true estimate to fall between if you redo the study many times.*

Reliability: *Similarly, **reliability** refers to the consistency of the measurement, or the ability to repeat the measurement and obtain the same result.*

National Library of Medicine



For example, a study reports a 95% confidence interval for a median exposure value. The true population median is 15. If the study were repeated many times with new samples, the calculated confidence intervals might vary (e.g., 12–18, 14–18, 14–16). However, with a 95% confidence level, about 95% of those intervals would still contain the true value (15).



Properties of Confidence Intervals (CI)

- **Fewer data lead to wider CIs**, increasing uncertainty
- **Poor-quality or highly variable data also widen CIs**
- **More data narrow the CI**, providing a more precise estimate
- **The width of the CI matters for risk assessment**, as large uncertainty may require more conservative decisions
- **Using the upper confidence limit (UCL) of a percentile** (e.g., 70th or 90th) adds a safety margin in risk assessment
- **CIs act as a safety net**, encouraging adequate data collection
- If an exposure estimate is **too close to a toxicological threshold value**, refinement may include collecting more data, narrowing the CI, or **improving data stratification** (e.g., defining better SEGs)



Content

1. Introduction
2. Understanding Percentiles & Distributions
3. Confidence Intervals & Data Requirements
- 4. Handling Left-Censored (Non-Detect) Data**
5. Software Tools for Exposure Data Analysis
6. Summary

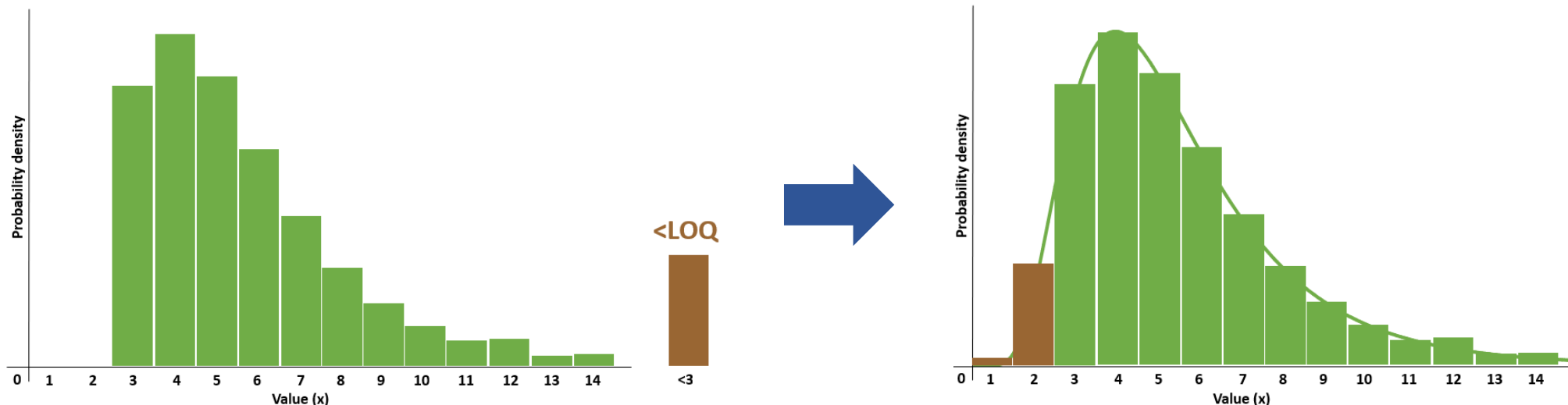
How to Handle Values below LOQ

Easy solution: Substitute non-detects with a value (e.g., 0, LOQ/2, LOQ/ $\sqrt{2}$, LOQ)

→ Simple but can bias results if many censored values

Advanced solutions: Statistical Estimation Methods

- Maximum Likelihood Estimation (MLE) or Regression on Order Statistics (ROS)
- Fit a distribution while accounting for censored data
- Provides more accurate percentile and confidence interval estimates





Content

1. Introduction
2. Understanding Percentiles & Distributions
3. Confidence Intervals & Data Requirements
4. Handling Left-Censored (Non-Detect) Data
- 5. Software Tools for Exposure Data Analysis**
6. Summary



Tools for Statistical Exposure Assessment

Purpose: Support risk assessors in analysing exposure distributions, defining SEG, handling censored data, and selecting appropriate percentiles

Examples of software tools (free to use):

- **BSOH – BWStat:** <https://www.bsoh.be/en/bwstatv310>
 - **Focus:** Worker risk assessment (EN689:2018)
 - Helps establish SEGs
 - Handles left-censored data
 - Calculates percentiles & confidence intervals according to EN689:2018
- **US EPA – ProUCL:** <https://www.epa.gov/land-research/proucl-software>
 - **Focus:** Environmental datasets (with/without censored data)
 - Handles left-censored data
 - Provides percentiles & confidence intervals for multiple distributions
 - Supports SEG establishment
- **R Foundation – R:** <https://www.r-project.org>
 - **Focus:** General statistical computing & graphics
 - Flexible tool for all discussed topics





Content

1. Introduction
2. Understanding Percentiles & Distributions
3. Confidence Intervals & Data Requirements
4. Handling Left-Censored (Non-Detect) Data
5. Software Tools for Exposure Data Analysis
6. **Summary**



Summary: Key Takeaways

- **Statistical approaches are essential** in risk assessment and must be understood by assessors
- **Percentiles reflect population exposure** and guide risk assessment decisions
- **Selecting the appropriate value from an exposure distribution is challenging** and depends on many factors, such as distribution type, population, toxicological endpoint, and data quality
- **Data reliability matters:** sample size, variability, and handling of left-censored data affect estimates
- **Software tools** can support distribution fitting, percentile estimation, and statistical analysis



Consequent Modules

Module 3:

- Introduction to Epidemiology

Future training videos:

- Additional legislations may be covered by ISES Europe in specialised training videos.

Thank You!



We appreciate your participation and attention

We encourage you to explore other ISES Europe training videos for deeper insights and broader understanding

Access all videos via: <https://ises-europe.org/>



Further Reading

Introduction to Statistics

Lane, DM

https://onlinestatbook.com/Online_Statistics_Education.pdf

Uncertainty and Data Quality in Exposure Assessment

World Health Organization (WHO) - International Programme on Chemical Safety (IPCS)

<https://www.who.int/publications/i/item/9789241563765>

The Importance of SEG Development

U.S. Environmental Protection Agency (EPA)

https://www.epa.gov/system/files/documents/2024-06/segs_arnold_sahmel.pdf