Exposure Assessment Statistics

Tips and Tricks

Peter Knott MClinEpid COH® MAIOH

NZOHS Conference, Auckland May 2019

Helping our clients keep their people safe





Outline of CES

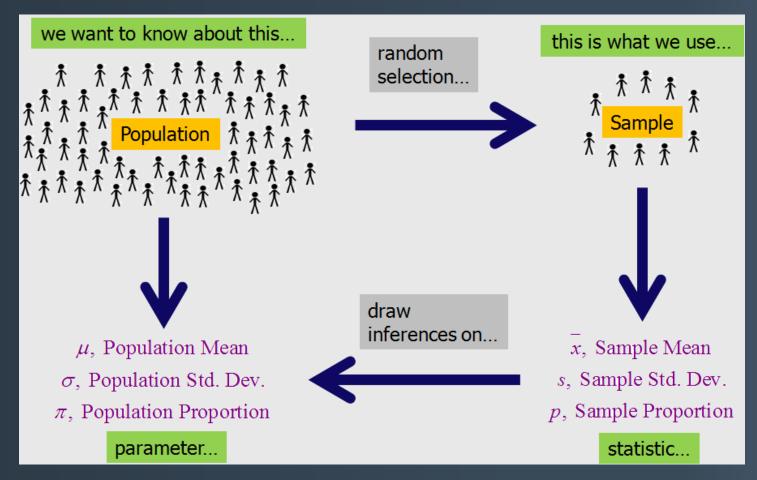
- Part 1
 - Sampling Strategies
 - Statistical methods for exposure data analysis
 - Examining the distribution of exposure data
 - Within- and between worker variance
 - What to do with censored data
 - Strategies for determining exceedance of a WES
 - Compiling Similar Exposure Groups (SEGs)
 - Assessing longitudinal results

- Part 2
 - AIHA's IHSTAT Excel[®]
 package
 - BOHS / NVvA BWStat
 package
 - Applications in proprietary statistical tools.
 - Expostats Bayesian toolkit for exposure assessment.

Why are you taking samples?

- Is exposure monitoring necessary?
- Obviously high exposures go straight to control focus
- Obviously insignificant exposures or uncertain?
- Use mathematical models to estimate / confirm instead
 - Advanced Reach Tool
- What is the purpose of the survey?
 - Compliance (Regulatory driven)
 - Baseline to derive Health Risk
 - Worst Case (Highest Exposed Worker) for control

Fundamentally...



Population v Sample

- A SEG with three workers would have a population of 780 possible 8-hour time-weighted average (8-hr TWA) exposures (52 weeks/year x 5 days/week x 3 people/day x 1 exposure/person = 780 exposures)
- The sample would consist of the 8-hr TWAs that were actually monitored (e.g. 10 full-shift samples)
- The estimate will be based on 1.3% of the exposures of the population
- If the SEG increases to 9 workers the population increases to 2340 full-shift exposures
- If the sample again contains 10 full-shift samples, the population estimate is now based on less than 0.5% of the full-shift exposures
- So given we are taking a small proportion you need to acknowledge uncertainty in your results

OK So I'm taking samples

- How much is enough?
- AIHA / NIOSH Approach (Leidel, Busch & Lynch 1977)
 - Probability of sampling highest risk employee in a population
 - More intense sampling effort
 - Not good if you have a small group of workers

TABLE A-1. SAMPLE S	SIZE FUR	K T	UP 1	10% (7=	= U.T) A	ND CUN	FIDENC	E 0.90	(a=0.1	, (001			
Size of group (N)	8	9	10	11-12	13-14	15-17	18-20	21-24	25-29	30-37	38-49	50	œ
Required No. of measured employees (n)	7	8	9	10	11	12	13	14	15	16	17	18	22

OK So I'm taking samples

• EN 689:2018 Approach

- Preliminary Test
 - Probability of obtaining exceedance based on lognormal distribution of exposures
 - Compares results obtained with a fraction of the WES for 3, 4 or 5 samples
- Limited sampling effort but works well for small groups of workers
- Additional sampling (at least 6 in total) needed for statistical test

Measurement Strategies – Worst Case

- Worst Case (Maximal Exposed Worker)
 - Historical Approach
 - Significant bias
 - No consideration of variability
 - How do you select the person?
 - May be useful for control assessment
 - Reduced sampling effort
- Still covered by:

s29 Ensuring prescribed exposure standards for substances hazardous to health are not exceeded

(1) A PCBU with management or control of a workplace must ensure that no person at the workplace is exposed to a substance hazardous to health in a concentration that exceeds the prescribed exposure standard for the substance.

• So a single exceedance is in breach of the Health and Safety at Work (General Risk and Workplace Management) Regulations 2016

Measurement Strategies – SEG Based

- Approach described in EN 689:2018
- Assembled after basic characterisation of workplace (s5.1)
- If estimates after basic characterisation >>WES = Sampling Control
- If estimates after basic characterisation <<WES = Sampling??
- If estimates after basic characterisation ??WES = Sampling
- 2 Stage Sampling Approach
- Randomly selected workers

Decide on your assessment and analysis protocol

- What you are going to sample for?
- How are you going to sample and analyse the results?
- Does the method you've chosen have sufficient performance and validation?
- What *is* your assessment strategy?
- What are your a priori SEGs and how did you pick them?
- What is an acceptable or unacceptable exposure assessment outcome?
- Write this down in a protocol (1 page) and then do your sampling

OK I've now got Data

Once you have a set of measurements look at:

- Simple Descriptive Info: min max
- Appropriate number of significant figures?
 - E.g. Respirable Dust reported to 2 sf & 1 decimal place for 5 place balance (eg 0.1, 1.2)
- How many Non-detects?
- Any weird results?
 - Lab / Math Errors

EN689 Preliminary Test

- No statistics!!
- Comparison to a lower limit of a small number of randomly selected samples
- 3 samples <0.1 x WES Threshold = Compliance
- 4 samples <0.15 x WES Threshold = Compliance
- 5 samples <0.2 x WES Threshold = Compliance
- Any result <WES or >Threshold = No Decision Additional Sampling
- Any result >WES = Non compliance Control

SEG	: Underground Closed Cab	in Drilling (8 employees in S	EG)
RCS Preliminary Results (mg/m ³)	Rp Preliminary Results (mg/m ³)	DPM Preliminary Results (µg/m3)	NO ₂ Preliminary Results (ppm)
0.009, 0.009, 0.003	0.35, 0.3, 0.2	83, 38, 55	<0.01, <0.01, 0.04
<10% WES? Yes	<10% WES? <mark>No</mark>	<10% WES? No	<10% WES? Yes
All Results <wes? Yes</wes? 	All Results <wes? Yes</wes? 	All Results <wes? Yes</wes? 	All Results <wes? Yes</wes?

Outcome of EN689 Preliminary Test

- 2nd Stage of Sampling for Rp and DPM
 - At least 3 more samples (min 6 in total)
- Then verification that less than 5% of exposures in SEG >WES (70% CI)

Efficiency of EN689 Preliminary Test

Agent	EN689 Sample No's.	Liedel 1977 Table A1 Sample No's.
RCS	3	7
Rp	6	7
DPM	6	7
NO ₂	3	7
Total	18	28

• 2 Drills

- Operated over 4 Crews
- 1 Operator / Drill
- 8 Employees in SEG
- More cost efficient strategy than NIOSH (Liedel et al 1977)

Periodic Sampling Strategy

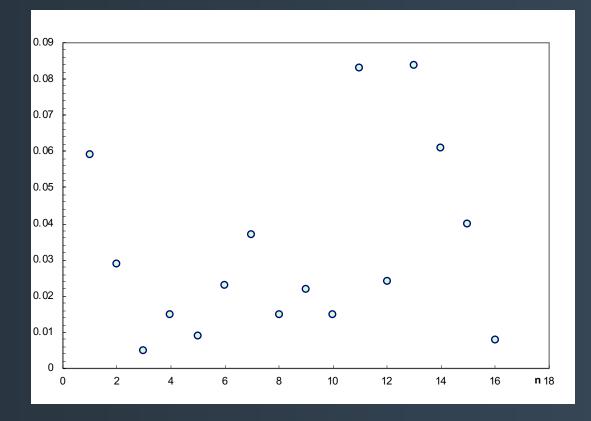
- Possibly the greyest area in sampling strategies
- Can range from complete reassessment of baseline / compliance to rarely if at all
- Driven by previous results and/or the presence / absence of any indicators of effect
- EN 689 contains options depending on fraction of WES
- AIOH Guidance on alternative methods

Elemental Carbon (mg/m ³)
0.059
0.029
0.005
0.015
0.009
0.023
0.037
0.015
0.022
0.015
0.083
0.024
0.084
0.061
0.040
0.008

Examining the distribution of exposure data

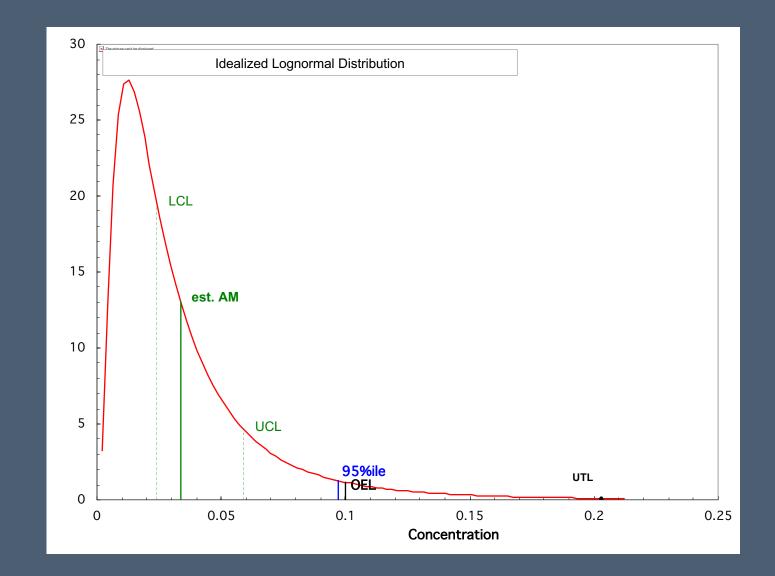
- 16 DPM TWA exposures at a site within a SEG taken over a 14 month period with no significant change in operations.
- OEL = 0.1mg/m^3
- No results >OEL
- GM = 0.024
- GSD = 2.31
- What do you think of this in terms of risk?
- What would you do if 4 samples had a sampling duration of 60% of the shift?

Examining the distribution of exposure data





Examining the distribution of exposure data



GSDs' in the real world – Workplace factors

LOW GSD	HIGH GSD
Highly controlled environments	Poorly controlled environments
Indoor operations	Outdoor operations
Jobs with one task per shift / week	Jobs with multiple tasks per shift / week
Background level of contamination masks effect of spatial and temporal variability	No background contamination to mask effect of spatial and temporal variability

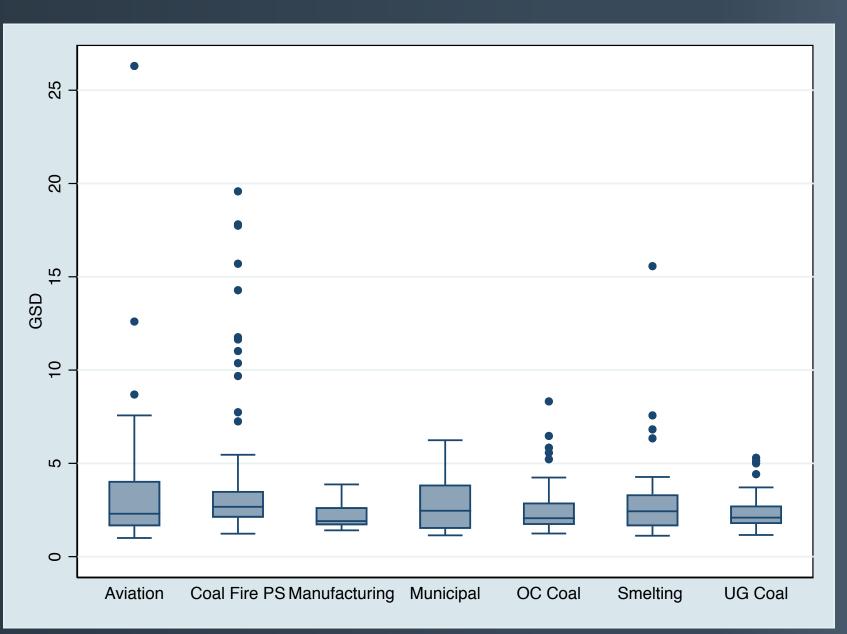
GSDs' in the real world – Assessment factors

LOW GSD	HIGH GSD
Insensitive measurements limited to small ranges (gravimetric, PCM)	Sensitive analytical measurements over many orders of magnitude (ICP, GC-MS)
One off or short temporal sampling periods (week)	Long term sampling periods (months / years)
Removal or substitution of non-detects with fixed value	Statistical treatment of non-detects with imputed values
Small sample sizes	Large sample sizes
Time weighted samples	Short duration samples (STEL, Real Time)

GSDs' in the real world

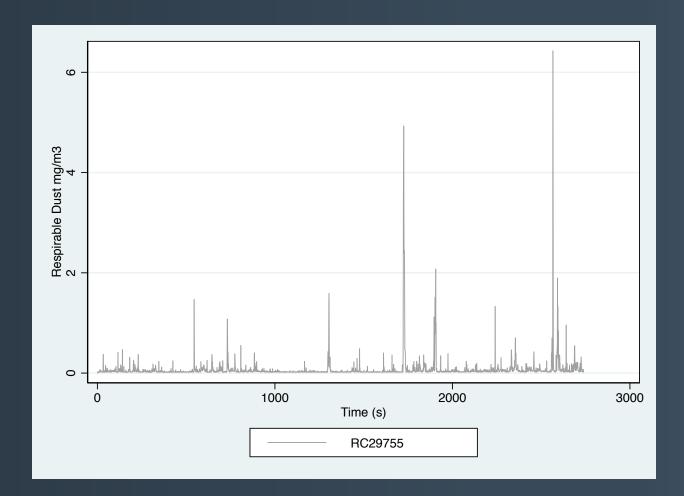
- Common statement of "GSD >[2.5, 3, 3.5] represent conditions where exposure variability is unacceptable and indicative of a poorly composed SEG or uncontrolled exposure"
- Inference is that results with GSD <2.5-3 are under control. [Workplace Factors] Without appreciation of assessment factors.
- Following slide represents approximately GSDs from 400 SEGs with different chemicals and facilities

25 20 15 . GSD 10 . ß • Assenic and soluble composition and a soluble composition of the solution of t Here C. The one times Nanage set and into provide the providence of the provide the providence of the prov

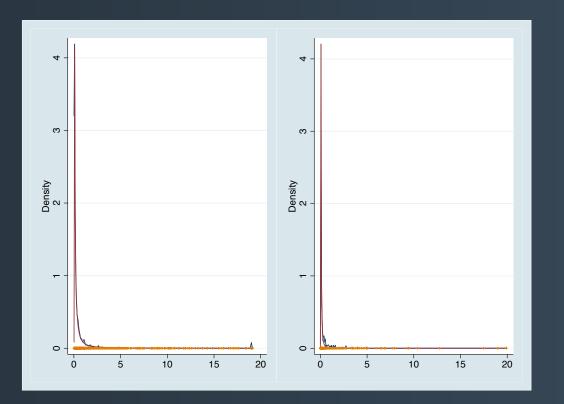




Real time Data

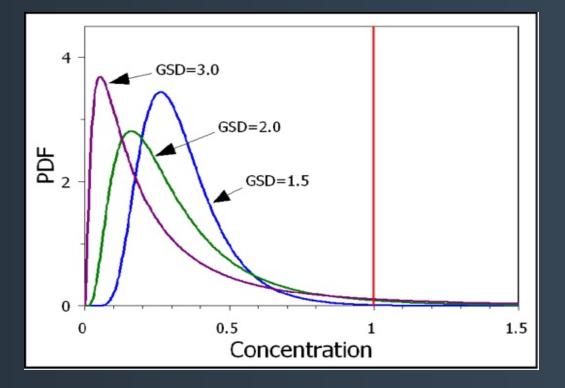


Extreme Value Distributions



- Characterised by high number of zero or low values and rapidly decreasing number of results away from the median
- Branch of statistics involved with the prediction of rare events (e.g.1/100year flood)
- Application to OH in frequency of peak exposures
- Not Lognormal so be careful if you are trying to describe or compare real time data.

Beware of increasing GSD on compliance



GSD	GM	MVUE	OEL	%EF
1.5	0.31	0.333	1	0.2%
2.0	0.26	0.333	1	2.7%
3.0	0.18	0.333	1	6.1%

Back to our EC data – all good?

DESCRIPTIVE STATISTICS	
Number of samples (n)	16
Maximum (max)	0.084
Minimum (min)	0.005
Range	0.079
Percent above OEL (%>OEL)	0.000
Mean	0.033
Median	0.024
Standard deviation (s)	0.026
Mean of logtransformed data (LN)	-3.713
Std. deviation of logtransformed data (LN)	0.839
Geometric mean (GM)	0.024
Geometric standard deviation (GSD)	2.315

LOGNORMAL PARAMETRIC STATISTICS	
Estimated Arithmetic Mean - MVUE	0.034
LCL _{1,95%} - Land's "Exact"	0.024
UCL _{1,95%} - Land's "Exact"	0.059
95th Percentile	0.097
UTL _{95%,95%}	0.203
Percent above OEL (%>OEL)	4.648
LCL _{1,95%} %>OEL	1.043
UCL _{1.95%} %>OEL	15.501

What about s29? No worker shall be exposed to a concentration >WES

What about Within and Between worker variance

- Concept originally developed in 1990's (Kromhout et al 1993)
- Individual Workers in similar groups are not similarly exposed
- Outdoor workers and those without LEV high exposure variability

Ann. occup. Hyg., Vol. 37, No. 3, pp 253-270, 1993 Printed in Great Britain. 0003-4878/93 \$6 00 + 0 00 Pergamon Press Ltd C 1993 British Occupational Hygiene Society.

A COMPREHENSIVE EVALUATION OF WITHIN- AND BETWEEN-WORKER COMPONENTS OF OCCUPATIONAL EXPOSURE TO CHEMICAL AGENTS

HANS KROMHOUT,*† ELAINE SYMANSKI† and STEPHEN M. RAPPAPORT†

*Department of Air Quality, Wageningen Agricultural University, P.O. Box 8129, 6700 EV Wageningen, The Netherlands; and †School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A.

(Received 31 December 1992 and in final form 23 February 1993)

Within and Between worker variance

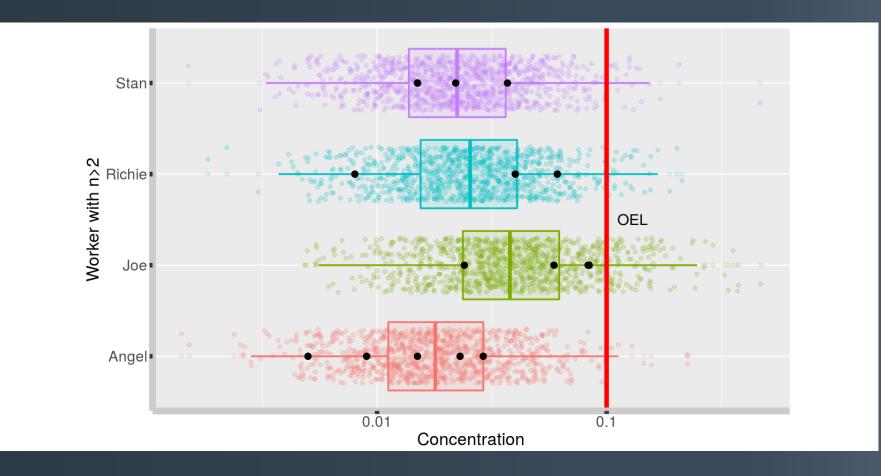
- Not fully addressed in AIHA strategy
- Central component of EN689
- Not widely used in Australia, but our EC data is an example
- Simple ANOVA of natural logs of results

Elemental Carbon (mg/m ³)	Worker
0.059	Joe
0.083	Joe
0.024	Joe
0.084	Joe
0.029	Angel
0.005	Angel
0.015	Angel
0.009	Angel
0.023	Angel
0.037	Stan
0.015	Stan
0.022	Stan
0.015	Stan
0.061	Richie
0.04	Richie
0.008	Richie

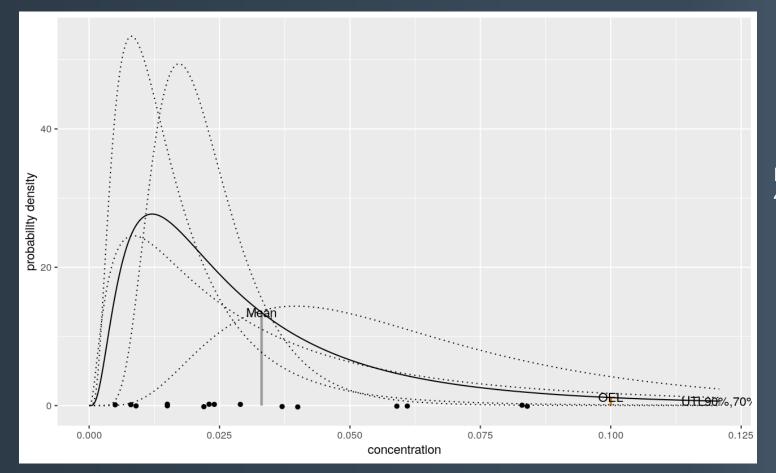
Within and Between worker variance

	Analysis	of Va	ariance		
Source	SS	df	MS	F	Prob > F
Between groups	4.6520688	3	1.5506896	3.14	0.0650
Within groups	5.91727762	12	.493106468		
Total	10.5693464	15	.704623094		
Bartlett's test for	equal varianc	es:	chi2(3) = 1.9	384 Prob	>chi2 = 0.

So our similar exposure group is not so similar



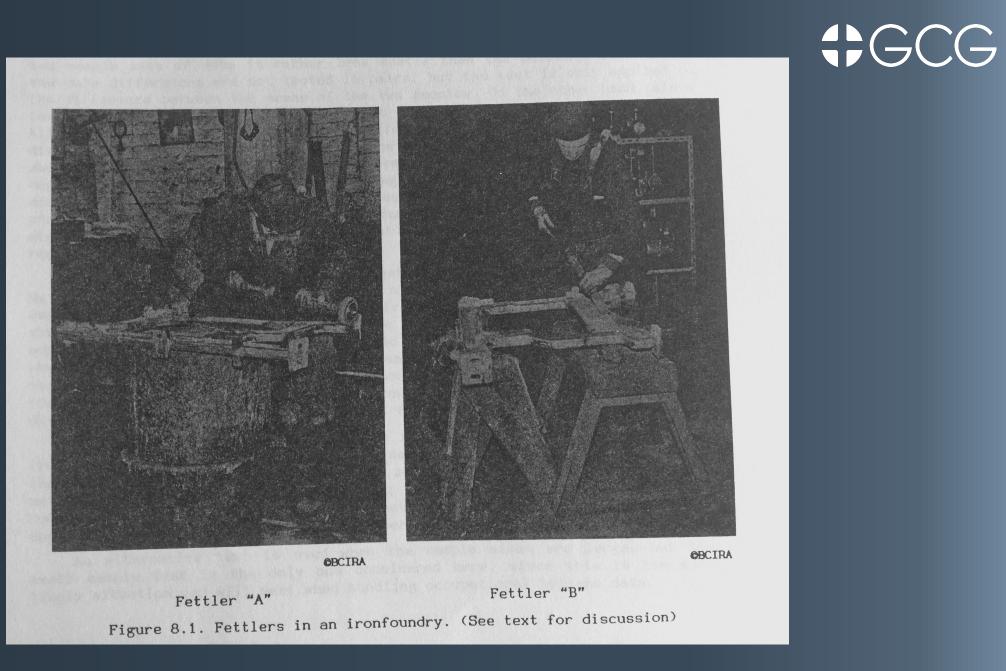
Evidence that there is non-compliance



Exposure profiles of our 4 workers

Historical Example from BOHS 1989

- Peter Dewell "Some Applications of Statistics in Occupational Hygiene"
 Fettler RCS Exposures
- Same tools, same castings, same number of castings
- Fettler A
 - Rp Exposure = 4.23mg/m3
 - RCS Exposure = 0.694mg/m3 (16% RCS)
- Fettler B
 - Rp Exposure = 1.45mg/m3
 - RCS Exposure = 0.192mg/m3 (13% RCS)



© Peter Dewell BOHS Tech Handbook Series No.1

Partial re-cap



Plan your assessment and how you will analyse the data – create an assessment protocol



Even though on a group basis the results may be suggestive of compliance individuals may be over exposed.



Don't rely on estimators of arithmetic mean or 95^{TH} percentile if you have small sample sizes (likely to have wide CI)



Exposure estimators of central tendency (GM, MVUE) will be affected by high GSD

What to do with censored (non-detect) data

Ann. Occup. Hys © The Author 2010. P on behalf of the B

Ann. Occup. Hyg., Vol. 51, No. 7, pp. 611-632, 2007 © The Author 2007. Published by Oxford University Press on behalf of the British Occupational Hygiene Society doi:10.1093/annhyg/mem045

Editorial

AUTHORS

Dave K. Verma

A Comparison of Several Methods for Analyzing Handling results below the level of detecti **Censored Data** T. L. OGDEN*

PAUL HEWETT¹* and GARY H. GANSER²

Chief Editor, Annals of Occupational Hygiene, British Occupational Hygiene Se Melbourne Court, Millennium Way, Derby DE24 8LZ, UK

AIHAJ 62:195–198 (2001)

Murray M. Finkelstein

Exposure Estimation in the Presence of Nondetectable Values: Another Look

249-263 2009 niversity Press ygiene Society nnhyg/men083

ised Imputation Approach for Data Analysis Ms. #173 esence of Non-detects

MOORTHY¹*, AVISHEK MALLICK¹ and THOMAS MATHEW²

athematics, University of Louisiana at Lafavette, Lafavette, LA 70504, USA: athematics and Statistics, University of Maryland, Baltimore, MD 21250, USA

Inn. Occup. Hyg., Vol. 54, No. 3, pp. 257-262, 201 © The Author 2009, Published by Oxford University Pres on behalf of the British Occupational Hygiene Society doi:10.1093/annhyg/mep092

Commentary

Much Ado About Next to Nothing: Incorporating Nondetects in Science **DENNIS HELSEL***

Practical Stats, 9278 Lark Sparrow Drive, Highlands Ranch, CO 80126, USA

#GCG

What to do with censored (non-detect) data

- Perennial Problem in IH and Environmental Monitoring
- Mostly wayyyyy too complicated for the average OH
- So most fall back on simple substitution LOD, LOD/2 LOD/ $\sqrt{2}$
- There is an alternative.....or two

#GCG

What to do with censored (non-detect) data

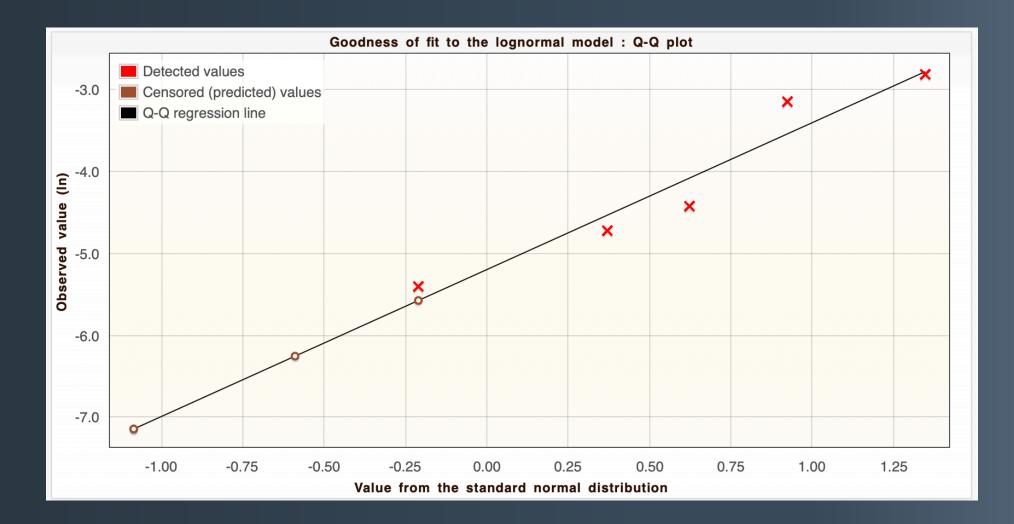
- EN689 recommendations: Do not substitute [fabricate an artificial distribution] of data
- Use of software to calculate regression coefficients for distributions containing censored values is recommended.
- NDexpo standalone application can impute data for inclusion in IH-STATS
- Expostats incorporates NDexpo
- BWStats incorporates NDexpo or EN689 (NADA)
- <u>http://practicalstats.com</u>

NDexpo

RCS mg/m3	Substitution (1/2 LOD)	
0.043	0.043	
<0.005	0.0025	
<0.005	0.0025	
<0.005	0.0025	
0.0089	0.0089	
0.012	0.012	
0.0045	0.0045	
<0.0036	0.0018	
0.06	0.06	

- Gets tricky when there is multiple censored values (<0.005, 0.0036)
- Regression methods deal with multiple, left and right censored data

Passer au français)		Switch to NADA mode Help file About	
NDExp	o – Treatme	ent of non-dete	tects in industrial hygiene samples	
Raw data	type detected	Final value 0.043	import export (final only) export (verbose)	
<0.005	ND	0.00079	0.043	
<0.005	ND	0.001923	0.0007901053035567333	
<0.005	ND	0.003794	_ 0.0019225039623145346 0.0037936156641240914	
0.0089	detected	0.0089	0.012 0.0045 0.007901053035567333	
0.012	detected	0.012		
0.0045	detected	0.0045		
<0.0036	ND	0.00079		
0.06	detected	0.06		
]			
New Rem		Calculate	Import	
dataset empty rows				



#GCG

• <u>NDexpo</u>

Determining exceedance of a WES

- 3 Parameters (not statistics) can be used:
- The 95% Upper confidence limit of the Minimum Variance Unbiased Estimator of Arithmetic mean (MVUE), aka UCL
- 95th Percentile of the exposure distribution (95%ile)
- Exceedance Fraction (%EF)
- Although in a strictly regulatory sense no instance of exposure above the WES is permitted

95% UCL of MVUE

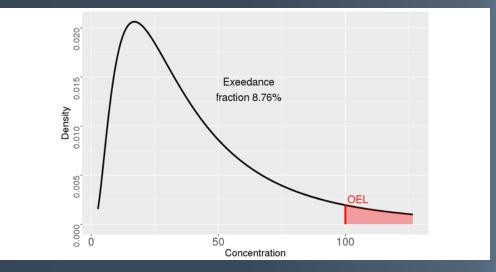
- The 95% UCL is an upper estimate of the minimum variance unbiased estimator of the population arithmetic mean. It is not specific to any OEL, and represents the upper limit of what the population arithmetic mean is likely to be.
- Compliance is then made to an OEL, the philosophy being if the 95%UCL is less than the OEL, then the population average is likely to be well below the OEL.
- Depends greatly on the Geometric Standard Deviation (GSD), a high GSD rapidly increases the UCL. High GSD's are prevalent in a number of situations as described.

95TH Percentile

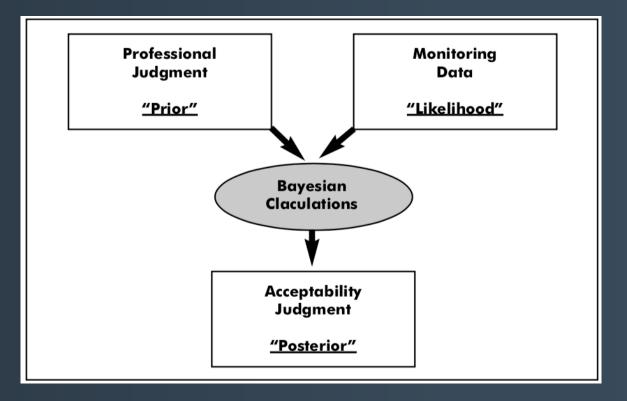
- The 95th percentile of the population exposures is the value which will be exceeded by no more than 5% of exposures
- A comparison is then made to the WES
- If 95th%ile <WES compliance likely

Exceedance Fraction

- A PDF is fitted to sample results and a population estimate obtained.
- The fraction of the distribution greater than an OEL is calculated. In this example the EEF is 8.76%, meaning that out of all the potential exposures in this SEG, 8.76% are likely to be >OEL
- Related to the 1977 NIOSH proposal that less than 5% of exposures should exceed the OEL.
- Comparing the exceedance fraction to 5% is numerically equivalent to comparing the estimated 95th percentile of the population distribution to the OEL.

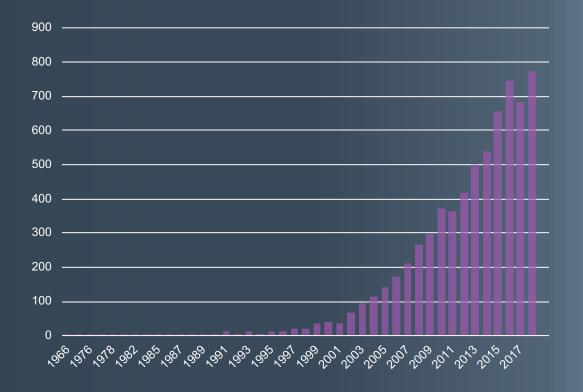


Bayesian Statistical Inference



Bayesian Statistical Inference

- Previously thought to be somewhat subjective as the Prior probability can come from non-quantitative sources
- BUT if you have quantitative data this can be used.
- Significant increase in use over past few years (PubMed data)
- Expressed in simple probability terms "There is X% probability that employee exposures are over the WES"



#GCG

Bayesian Statistical Inference

- Applied by AIHA "A strategy for assessing and managing occupational exposures" text
- Exposure Categories:
 - 0 (<1% of OEL)
 - 1 (<10% of OEL)
 - 2 (10-50% of OEL)
 - 3 (50-100% of OEL)
 - 4 (>100% of OEL)
- Used by Expostats with use of weakly informative prior for estimates of μ and informative prior for σ based on historical estimates GSD



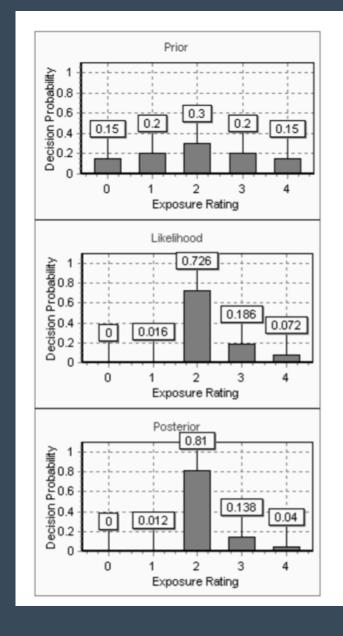
How does it work



Results

If you have no prior data

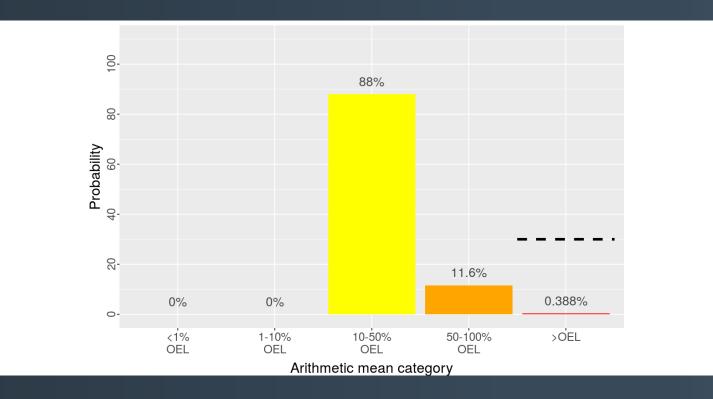
Non-informative or Weakly informative Priors



©AIHA 2008

Our EC data

• How likely is it that my exposure parameter falls in a certain interval?



Compiling / Testing Similar Exposure Groups (SEGs)



AIHA Definition



"A group of workers having the same general exposure profile for an agent because of the similarity and frequency of the tasks they perform, the similarity of the materials and processes with which they work, and the similarity of the way that they perform the tasks."

How do I compile a SEG

- Begin with your HRA characterisations
- Look at your agent inventory
- Don't forget components, intermediates, by- products, maintenance, physical/biological agents
- Divide the workplace into processes, then jobs/tasks, and look for similarities
- Ask about infrequent but routine tasks
- STELs or Ceiling Limits may mean task-based SEGs
- Be careful about creating too many SEGs simple is often the best

#GCG

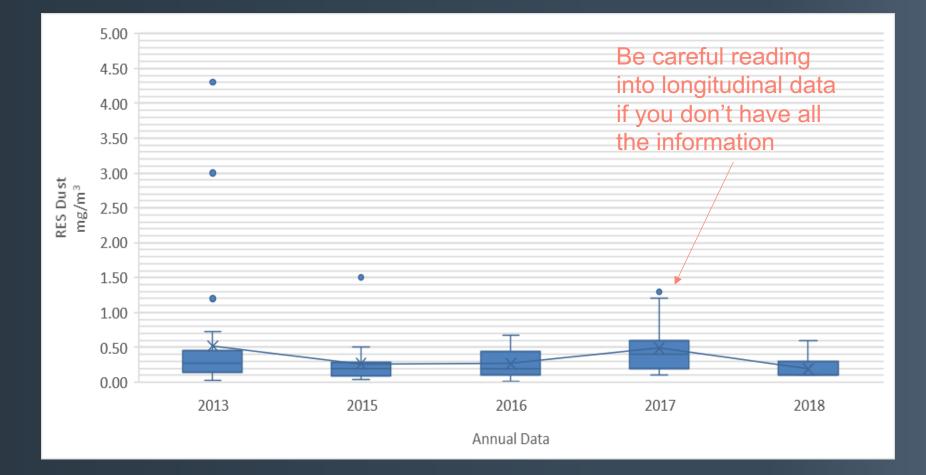
How do I compile a SEG

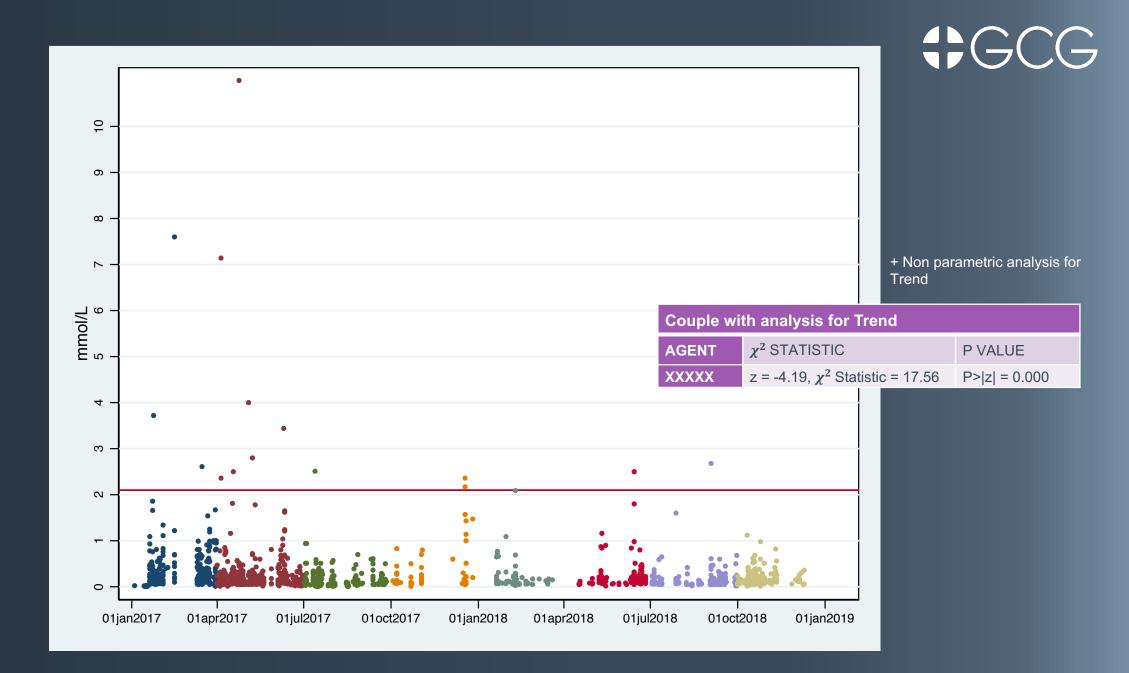
- Observational (qualitative)
 - Observe activities in the workplace
 - Make judgments on expected similarities in exposures to various agents
 - Employ mechanistic models (Advanced Reach Tool?)
- Sampling (quantitative)
 - Measure exposures
 - Apply statistical analysis
- Most of the time you won't be able to make a quantitative judgement
- Quantitative Assessment of SEGs requires data (often a lot)

Methods for assessing longitudinal results

- Boxplots
- Scatterplots
- Control charts
- Trend analysis
- Non-parametric regression
- Bayesian Decision Analysis

Boxplots by Year



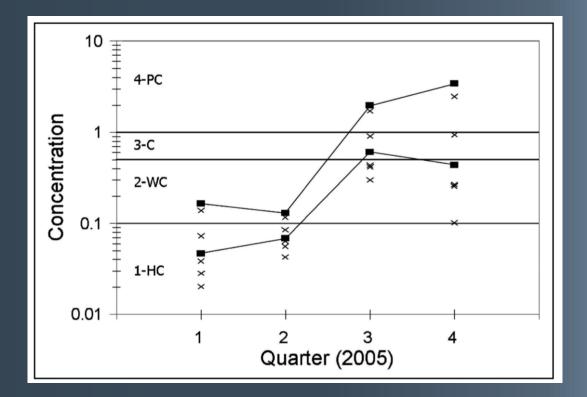




GCG

Control Chart Approach

- AIHA Suggestions
 - Plot sequential samples against action and warning limits
 - Plot running 95th Percentile
 - Plot GM and running 95th Percentile



Bayesian Interval Probabilities



AIHA IHSTAT Package

Protectina Worker Health

If this file doesn't work....

Enable macros when opening this file.

Attivare le macro all'apertura del file

Ativar macros guando abrir este arguivo.

Aktiver makroer når du åpner denne filen

इस फ़ाइल खोलने जब स्थूल सक्षम है.

이 파일을 열 때 매크로를 활성화함.

ファイルを開く時マクロを有効にしてください

Multilingual IHSTAT+

The material embodied on this software is provided "as-is" and without warranty of any kind, expressed, implied or otherwise, including without limitation any warranty of merchantability or fitness for a particular purpose.

In no event shall John R. Mulhausen, Ph.D., CIH, or the American Industrial Hygiene Association (AIHA) be liable for any direct, indirect, special, incidental, or consequential damages of any kind, or any damages whatsoever, including without limitation loss of profit, loss of use, savings or revenue, or the claims of third parties, whether or not John Mulhausen or the AIHA has been advised of the possibility of such loss. however caused, and on any theory of liability, arising out of or in connection with the possession, use, or performance of this software.

> This file was originaly created by John Mulhausen and then modified in its multilingual version by Daniel Drolet et al.

Habilite los macros cuando abra este archivo.

Při otevření tohoto souboru povolte makra.

Dosvavı acarken makroları etkinlestirin

Macro's inschakelen bij het openen van dit bestand

Запускайте работу макросов при открывании документа

注意:打开该文档时请启用宏。





2007

2000 / 2003

2010

2013

A full discussion on how to analyze and interpret exposure monitoring data can be found in the publication



Ignacio, J. and, Bullock, B. (editors) A Strategy for Assessing and Managing Occupational Exposures, 3rd Edition. Fairfax, VA: AIHA Press, 2006

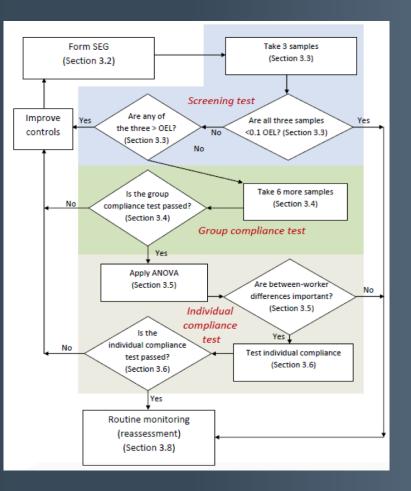
IHSTAT+ : v. 235, Dec 2013

This file requires that macro security level of Microsoft Excel must be set in order to enable MACROS. For more information, refer to the Microsoft Web site:

Activer les macros à l'ouverture du fichier. Beim Öffnen der Datei Makros aktivieren.

BOHS / NVvA BWStat Package

- Excel package
- Web Hosted interface





R

Stata

Stata

IHDataAnalyst

Flexible Bayesian Application

Specifically designed for OH data (Incorporated into Enterprise system) Uses AIHA Exposure Interval Classes (<10%, 10-50%, 50-100%, >100% User defined Prior probabilities or generic priors Fixed Graphical outputs

Commercial Statistics Package Extensive Analysis Options with programming ability **Bayesian Options** Flexible graphics outputs

Powerful Open source statistical package Command line language IH packages available **Extensive Bayesian Options** Flexible graphics outputs

Proprietary Tools

#GCG

Expostats

Annals of Work Exposures and Health, 2018, 1–13 doi: 10.1093/annweh/wxy100 Original Article

> The Chartered Society for Worker Health Protection

OXFORD

Original Article

Expostats: A Bayesian Toolkit to Aid the Interpretation of Occupational Exposure Measurements

Jérôme Lavoué^{1,2,*}, Lawrence Joseph³, Peter Knott⁴, Hugh Davies⁵, France Labrèche^{1,6}, Frédéric Clerc⁷, Gautier Mater⁷ and Tracy Kirkham⁸

Expostats

http://expostats.ca/site/en/index.html

So....

- Sampling Strategies
- Statistical methods for exposure data analysis
- Examining the distribution of exposure data
- Within- and between worker variance
- What to do with censored data
- Strategies for determining exceedance of a WES
- Compiling Similar Exposure Groups (SEGs)
- Assessing longitudinal results
- AIHA's IHSTAT Excel® package
- BOHS / NVvA BWStat package
- Applications in proprietary statistical tools.
- Expostats Bayesian toolkit for exposure assessment.





Thank you for your patience, participation and presence.

peter.knott@gcg.net.au