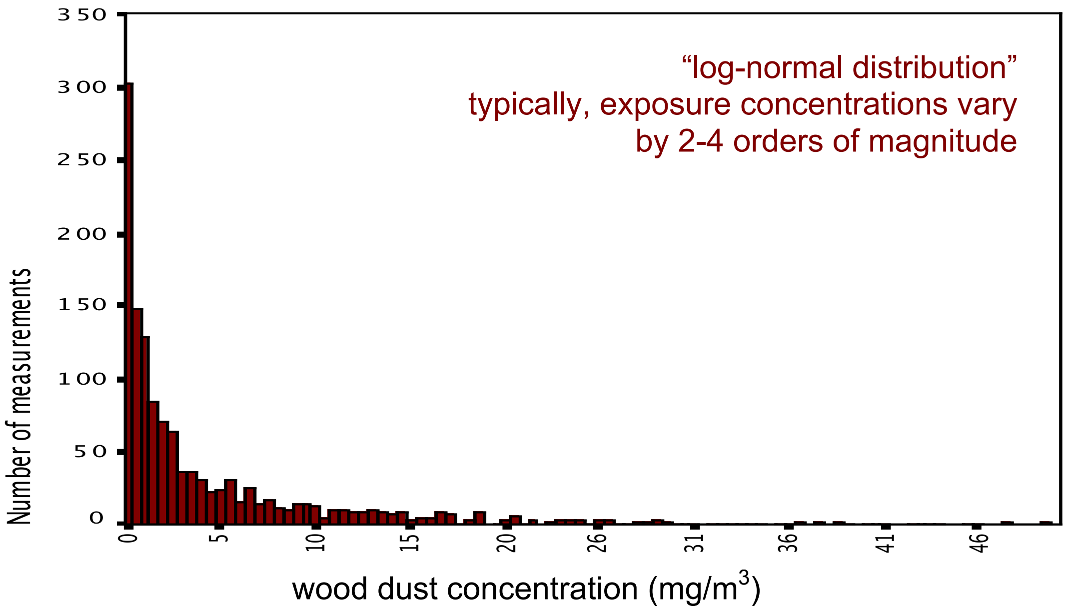
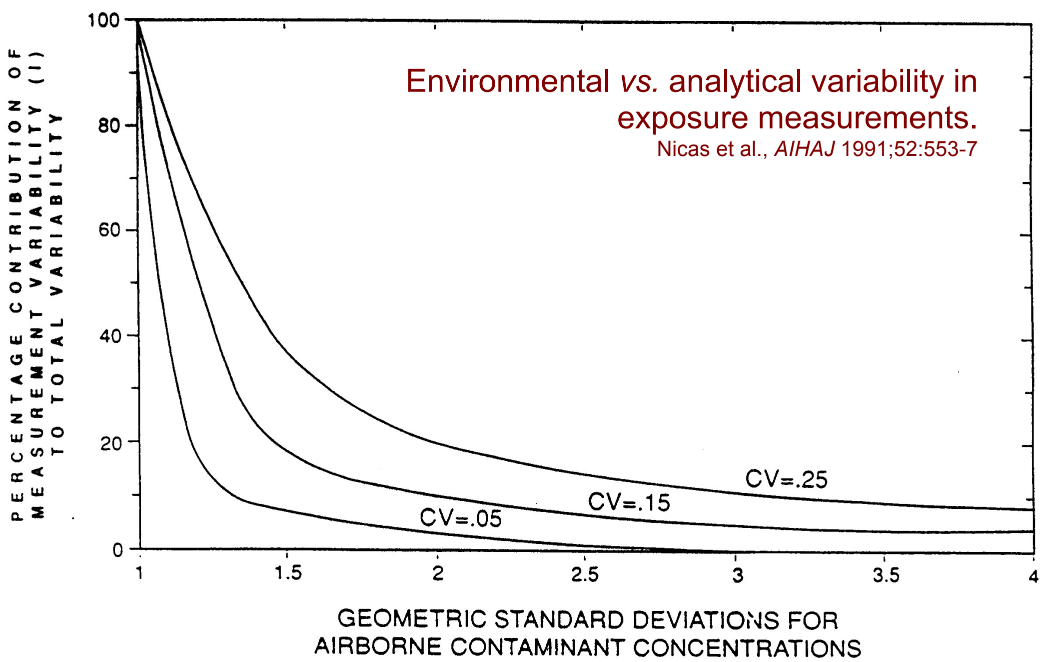
SAMPLING STRATEGIES

What is the importance of sampling strategy?



Measurement error is usually a small proportion of environmental variability:



ELEMENTS OF SAMPLING STRATEGY

Temporal components

1. How long to measure
2. When to measure
3. How often to measure

Spatial components

4. Where to measure
5. Who to measure

Statistical components

6. Method of sample selection
7. Number of samples
8. Statistical analyses
9. Supplementary data collection

SOME PURPOSES OF EXPOSURE MEASUREMENT

A. **Compliance**, comparisons with standards
   - continuous monitoring of acute hazards
     - confined spaces
   - periodic monitoring of chronic hazards

B. **Determinants of exposure**
   - identifying factors influencing exposure levels
     - locations
     - tasks
     - processes
     - control measures
     - equipment
     - environmental conditions
     - personal characteristics

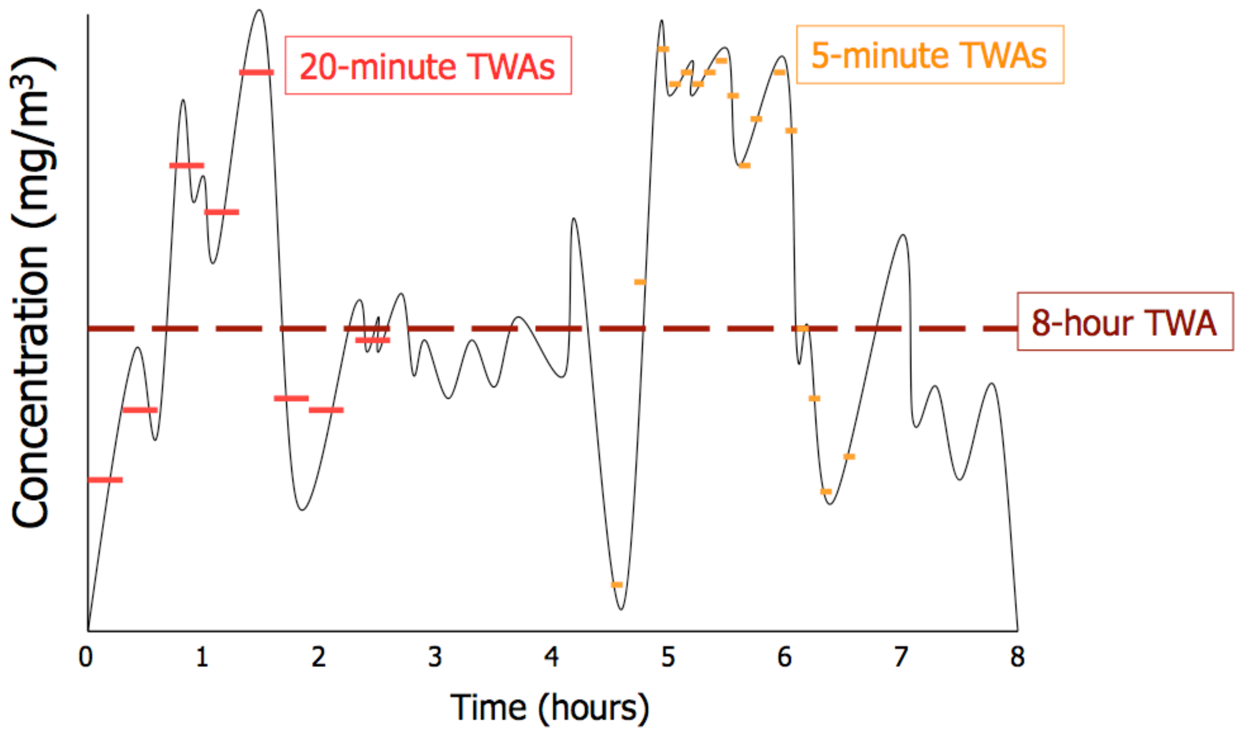
C. **Epidemiologic studies**
   - establishing exposure-response relationships

D. Evaluating measurement methods
   - validity
   - reliability

E. Risk assessment
   - understanding exposure levels across populations

Temporal Issues: How Long to Measure

Example of air concentrations over 8 hour period, e.g., typical work shift:



- longer averaging times remove peaks and valleys in concentrations

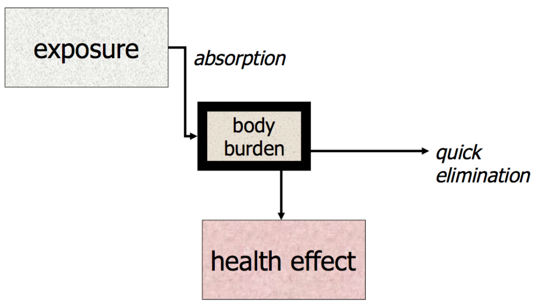
- all measuring devices average to some extent:
  - filtering devices or passive monitors average over entire measuring period (called moving time averagers; sample mass increases with increased sampling time)
  - direct reading instruments have a given response time over which averaging takes place (called exponential averagers; like the body, always pulling sample in and sending it out)

How should we decide on a sampling duration?

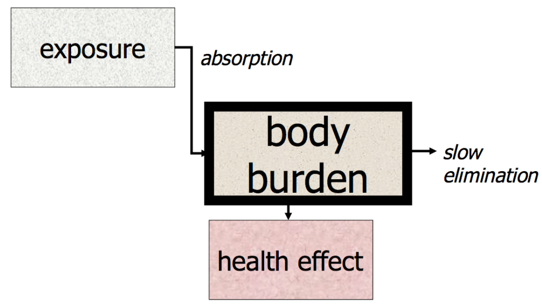
- relate sampling duration to time variations in body burden, which depend on the biological half-life of the chemical ($T_{½}$)
  - using this idea, Roach 1966, 1977 came up with the following sampling durations:

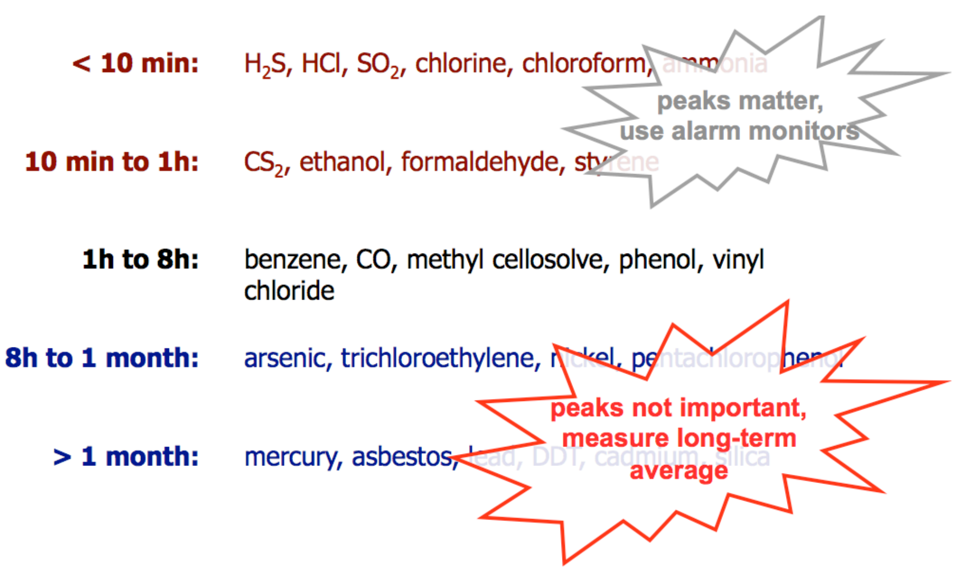
- if the duration of measurement is 2.88 * $T_{1/2}$ then the variability in the exposure measurement would equal the variability in the body
- sometimes need a safety factor, i.e., if the effect is acute and severe; Roach suggests sampling times of 1/10 to 1/2 * $T_{1/2}$ should be short enough for all conceivable situations

Short $T_{1/2}$



Long $T_{1/2}$





SPATIAL ISSUES: PERSONAL *vs.* AREA MEASUREMENTS

Data from a nuclear power plant in Britain, reported in the *Annals of Occupational Hygiene*, 1969;12:33-40, by D.C. Stevens:

Location	**Operator**	**Ratio of Exposure as Indicated by Personal Monitor to that Indicated by Area Monitor**	
		Mean Alpha Activity	**Mean Beta Activity**
Incinerator	A	1.4	3.0
	B	0.49	2.1
	C	0.49	2.4
Filter House	D	1.9	2.7
Decontamination Area	E	4.1	5.5
	F	8.8	18.0
	G	20.5	20.0
	H	7.1	17.0
	I	10.2	12.3
	J	11.3	20.8

Data from an aluminum smelter Soderberg potroom, reported in *IARC Monographs* on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Volume 34, 1984, p. 45:

Substance	**Mean of Personal Monitors**	**Mean of Area Monitors**	**Ratio P:A**
Hydrogen fluoride	1.95 mg/m^3	0.34 mg/m^3	5.73
Alumina	4.05	3.50	1.16
Pitch volatiles	18.0	0.57	31.6
Benzo-a-pyrene	37.0	2.78	13.3

Data about inhalable particulate concentrations in magnesium and aluminum productions facilities (a foundry and 3 smelters) in Quebec, reported in the *Journal of Occupational and Environmental Hygiene*, 2009;6:687-697 by Dufresne *et al*:

Area	**Median of Personal Monitors**	**Median of Area Monitors**	**Ratio P:A**
1	2.24 mg/m^3	0.78 mg/m^3	2.9
2	34.0	2.40	14.2
3	9.80	6.10	1.6
4	4.00	6.00	0.7
5	23.0	1.05	22.0
6	4.75	3.10	1.5

STATISTICAL ISSUES: WORST CASE VS. RANDOM SAMPLING

- this issue applies to considerations of **who** and **when** to measure

- measurements of **"worst case"**
  - can be useful as a way to screen exposures, especially for compliance monitoring
  - idea is that if exposures in worst case time or place are low, then don't need to sample in other times or places
  - major concern: that your judgment about what is worst case is incorrect
  - also concern that exposure estimates are biased to the high side, therefore data not useful for other purposes, such as epidemiology or determinants modeling

- measurements of **random sample of population**
  - allows statistical inferences
  - does not require "professional judgment", which can be prone to error
  - gives a picture of exposures over widely varying times, people
  - can take "simple" random sample (list all possible locations, people, or days, select randomly from list
  - can also take "stratified" random sample (e.g., group locations or people with similar exposure potential, then take random sample from each group)
  - useful not only for compliance sampling, but also for epidemiology, determinants of exposure modeling
  - most spreadsheet and statistical programs have random number generators

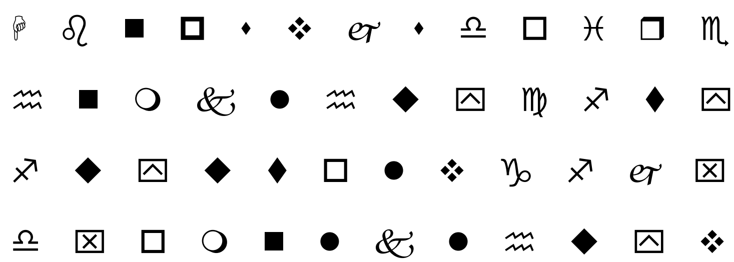
- measurements of **whole population**
  - in certain cases, e.g., quick-acting severely toxic agents, **all** members of the potentially exposed population need to be monitored
    - e.g., personal alarm monitors in confined spaces

Statistical Issues: Methods of Sample Selection

Random Sampling Methods:

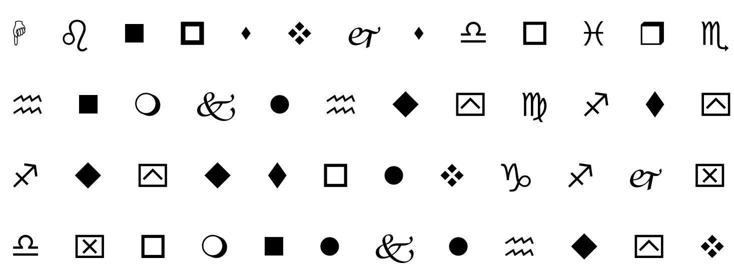
Simple Random Sampling

Simple random sampling is the basic sampling technique used to select a sample from a larger group (a population). Each day and time is chosen entirely by chance and each member of the population has an equal chance of being included in the sample. It is easy to set up, but may be difficult logistically to carry out, and with small samples, may not get even coverage of characteristics of interest. Use a table of random numbers, a computer random number generator, or a mechanical device to select the sample.



Systematic Sampling with a Random Start

Taking measurements at set intervals, e.g., every 5th day, or every mile along a street, but select the initial sample at random. This method can cause systematic biases, *if* the systematic pattern selected is related to patterns in the data (e.g., if you randomly selected Monday as the starting day of your sampling and then sampled every 7 days, and exposures on Mondays were different than on other days of the week).

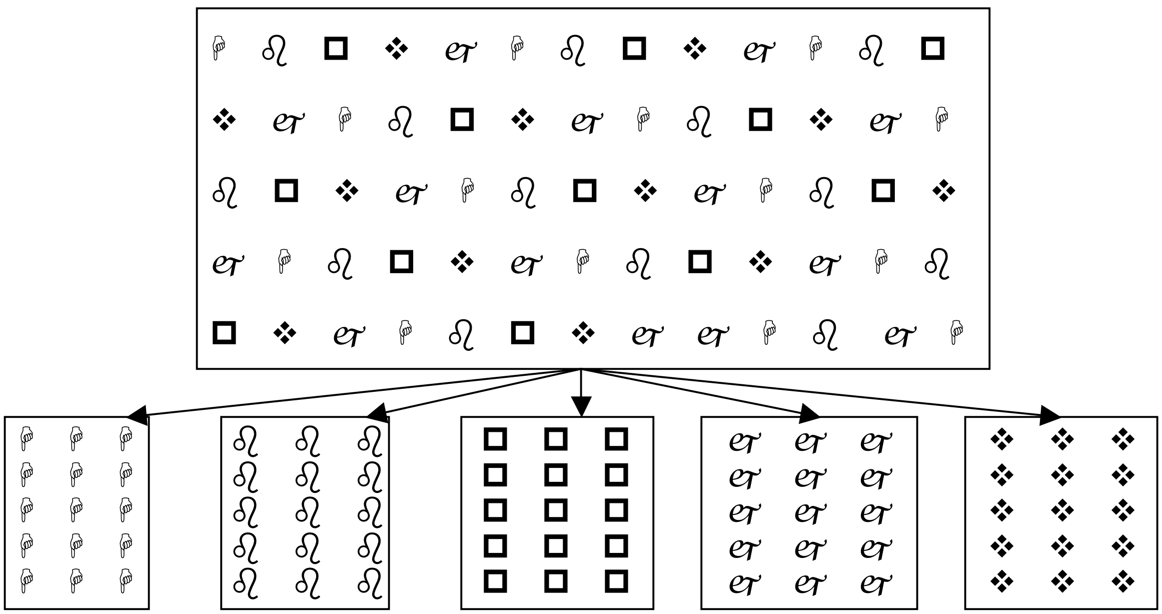


Stratified Random Sampling

There may be factors which divide the population into sub-populations (strata) and we may expect the measurement of interest to vary among the different sub-populations (e.g., job, sex, age, height of employees in occupational exposure measurement; proximity to source, city vs. rural residence, age, sex of subjects in environmental exposure measurement).

To make sure that each sub-group is adequately represented in the sample, first identify the members of each stratum, then randomly sample from each. Can sample so that

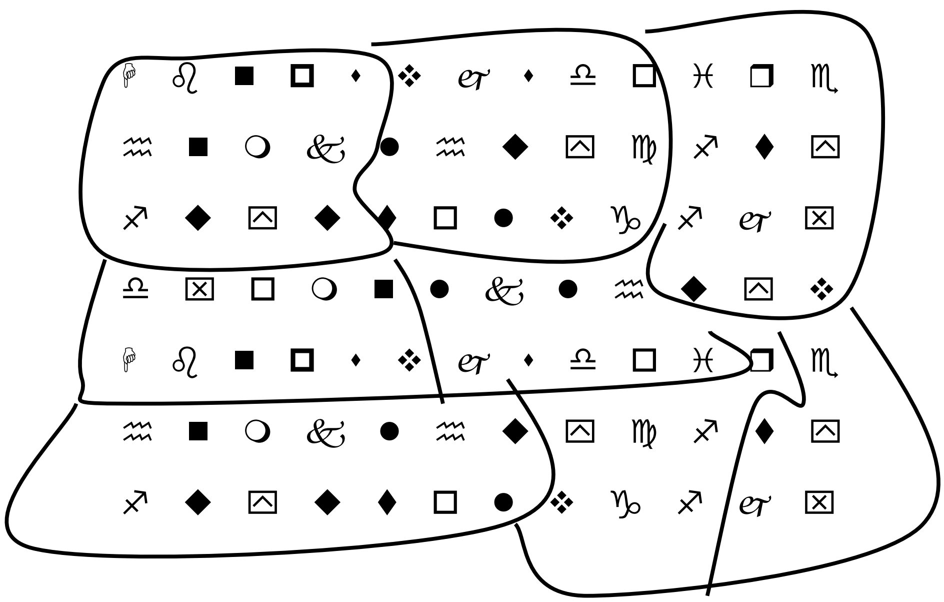
- the proportion of each stratum in the sample is the same as in the population, or
- certain strata that have fewer members or greater exposure variability are oversampled, to ensure enough samples for statistical inferences



Cluster Sampling

Cluster sampling is a sampling technique whereby the entire population is divided into segments (usually geographic), i.e., "clusters", and a random sample of these clusters is selected. In true cluster sampling, you would then measure all exposures within the randomly selected clusters. However, often there is then random sampling of the space/time element of interest within the cluster (this is then called "multistage sampling," i.e., sampling using multiple techniques).

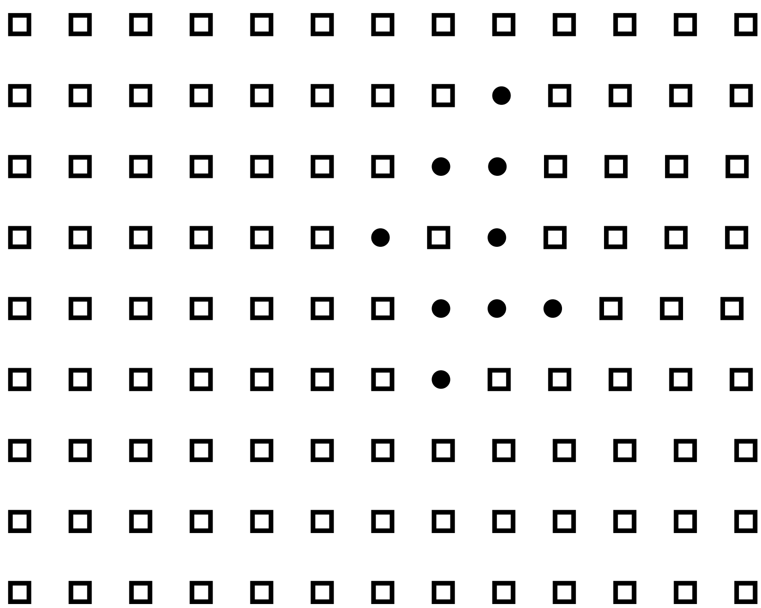
Cluster sampling is typically used when it is difficult to get a complete list of the members of a population, but a list of clusters is possible. It is also used when a random sample would produce a list of subjects so widely scattered that surveying them would prove to be far too expensive. There are special methods of analysis with this kind of sampling to prevent biased results.



Adaptive Sampling

Adaptive sampling was conceived as a response to the problem of sampling rare populations that are likely to be close to each other in space. Examples might include contamination from a point source or endangered species that live in groups. This kind of sampling involves an initial random sample and whenever the variable of interest satisfies a condition, additional samples are taken near the sample of interest. There are special methods of analysis with this kind of sampling to prevent biased results.

This is sometimes called "adaptive cluster sampling," but it is not related to cluster sampling above. Here the term "cluster" refers to geographic clustering of rare items.



Non-Random Methods:

Convenience Sampling

Taking measurements at a time and place that is feasible for the person doing the measurements. The sample is not a random sample and therefore the sampling distributions of any statistical parameters are unknown.

Quota Sampling

Quota sampling is a method of sampling widely used in opinion polling and market research. Interviewers are each given a quota of subjects of specified type to attempt to recruit. For example, an interviewer might be told to go out and select 20 adult men and 20 adult women, 10 teenage girls and 10 teenage boys for exposure measurements. It suffers from a number of methodological flaws, the most basic of which is that the sample is not a random sample and therefore the sampling distributions of any statistical parameters are unknown.

Some Classic and Summary Documents on Log-Normal Distributions, Sampling Strategies and Compliance Decision Making

---

Aitchison J, Brown JAC. *The Lognormal Distribution.* Cambridge University Press, 1957

Ashford JR. The design of a long-term sampling programme to measure the hazard associated with an industrial environment. *J Royal Stat Soc* 1958;121:331-47

Boleij JSM, Buringh E, Heederik D, Kromhout H. Monitoring strategies for compliance. In *Occupational Hygiene of Chemical and Biological Agents.* Amsterdam: Elsevier, 1995, p. 147-156

British Occupational Hygiene Society. *Sampling Strategies for Airborne Contaminants in the Workplace.* BOHS Technical Guide No. 11, H and H Scientific Consultants, Leeds, UK, 1993

British Occupational Hygiene Society & Nederlandse Vereniging Arbeidshygiene. *Testing Compliance with Occupational Exposure Limits for Airborne Substances.* 2011

Brunekreefe B, Noy D, Clausing P. Variability of exposure measurements in environmental epidemiology. *Am J Epidemiol* 1987;125:892-898

Burstyn I, Teschke K. Studying the determinants of exposure: A review of methods. *Am Ind Hyg Assoc J* 1999;60:57-72

CEN. Workplace atmospheres – Guidance for the assessment of exposure to chemical agents for comparison with limit values and measurement strategy. prEN 689. 1992

Corn M. Strategies of air sampling. *Scand J Work Environ Health* 1985;11:173-180

Correspondence between Paul Hewitt and Steve Rappaport, *Annals of Occupational Hygiene* 1998; 42(6):413-422

Friesen MC, MacNab YC, Marion SA, Davies HW, Teschke K, Demers PA. Mixed models and empirical Bayes estimation for retrospective exposure assessment of dust exposures in Canadian sawmills. *Ann Occup Hyg* 2006;50:281-288.

Hawkins NC, Norwood SK, Rock JC. *A Strategy for Occupational Exposure Assessment.* Akron: American Industrial Hygiene Association. 1991

Hewett P. Misinterpretation and misuse of exposure limits. *Appl Occup Environ Hyg.* 2001;16(2):251-6

Keith LH. *Principles of Environmental Sampling* (2nd edition). Oxford University Press: Oxford, 1996

Kromhout H, Loomis DP, Kleckner RC, Savitz DA. Sensitivity of the relation between cumulative magnetic field exposure and brain cancer mortality to choice of monitoring data grouping scheme. *Epidemiol* 1997;8(4):442-445

Kromhout H, Loomis DP, Mihlan GJ, Peipins LA, Kleckner RC, Iriye R, Savitz DA. Assessment and grouping of occupational magnetic field exposure in five electric utility companies. *Scand J Work Environ Health.* 1995;21:43-50

Leidel A, Busch KA, and Lynch JR. *Occupational Exposure Sampling Manual.* US Dept of Health, Education, and Welfare, NIOSH, 1977

Mackay D, Paterson S. Spatial concentration distributions. *Environ Sci Technol* 1984;18:207A-214A

McCracken JP, Schwartz J, Bruce N, Mittleman M, Ryan LM, Smith KR. Combining individual- and group-level exposure information. Child carbon monoxide in the Guatemala woodstove randomized control trial. *Epidemiology* 2009; 20:127-136.

Mikkelsen AB, Schlunssen V, Sigsgaard T et al. Determinants of wood dust exposure in the Danish furniture industry. Ann Occup Hyg 46:673-85. 2002.

Mulhausen JR, Damiano J. *A Strategy for Assessing and Managing Occupational Exposures*, 2nd Edition. Fairfax,VA: AIHA Press. 1998

Neitzel RL, Daniell WE, Sheppard L, Davies HW, Seixas NS. Improving exposure estimates by combining exposure information. Ann Occup Hyg 2011; doi: 10.1093/annhyg/mer011.

Nicas M, Simmons BP, Spear RC. Environmental versus analytical variability in exposure measurements. *Am Ind Hyg Assoc J* 1991;52:553-7

Rappaport SM, Selvin S. A method for evaluating the mean exposure from a lognormal distribution. *Am Ind Hyg Assoc J* 1987;48:374-379

Rappaport SM. Assessment of long-term exposures to toxic substances in air. *Ann Occup Hyg*, 1991;35:61-121

Rappaport SM. Smoothing of exposure variability at the receptor: Implications for health standards. *Ann Occup Hyg* 1985;29:201-214

Rappaport SM. Lyles RH. Kupper LL. An exposure assessment strategy accounting for within- and between-worker sources of variability. *Annals of Occupational Hygiene*. 1995:39(4):469-95

Roach SA. A more rational basis for air sampling programs. *Am Ind Hyg Assoc J* 1966;27: 1-12.

Roach SA. A most rational basis for air sampling programs *Ann Occup Hyg*. 1977;20:65-84.

Saltzman BE. Significance of sampling time to air monitoring. *J Air Poll Control Assoc* 1970;20:660-665

Seixas NS, Sheppard L. Maximizing accuracy and precision using individual and grouped exposure assessments. *Scand J Work Environ Health* 1996;22:94-101

Singh A, Singh AK, Maichle RW. *ProUCL Version 3.0 User Guide*. US Environmental Protection Agency, 2004.

Symanski E, Greeson NMH, Chan W. Evaluating measurement error in estimates of worker exposure assessed in parallel by personal and biological monitoring. *Am J Ind Med* 2007; 50:112-121.

Tielemans E, Kupper LL, Kromhout H, Heederik D, Houba R. Individual-based and group-based occupational exposure assessment: some equations to evaluate different strategies. *Annals Occup Hyg* 1998;42(2):115-9

Weaver MA, Kupper LL, Taylor D, Kromhout H, Susi P, Rappaport SM. Simultaneous assessment of occupational exposures from multiple worker groups. *Annals of Occupational Hygiene* 2001;45(7):525-42.