Institut de recherche
Robert-Sauvé en santé
et en sécurité du travail

WebExpo
Towards a Better Interpretation of Measurements of Occupational Exposure to Chemicals in the Workplace

Jérôme Lavoué
Lawrence Joseph
Tracy L. Kirkham
France Labrèche
Gautier Mater
Frédéric Clerc

STUDIES AND RESEARCH PROJECTS

R-1065

CHEMICAL AND BIOLOGICAL HAZARD PREVENTION RESEARCH

OUR RESEARCH
is working for you !

The Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST), established in Québec since 1980, is a scientific research organization well-known for the quality of its work and the expertise of its personnel.

Mission

To contribute, through research, to the prevention of industrial accidents and occupational diseases and to the rehabilitation of affected workers;

To disseminate knowledge and serve as a scientific reference centre and expert;

To provide the laboratory services and expertise required to support the public occupational health and safety network.

Funded by the Commission des normes, de l'équité, de la santé et de la sécurité du travail, the IRSST has a board of directors made up of an equal number of employer and worker representatives.

To find out more

Visit our Web site for complete up-to-date information about the IRSST. All our publications can be downloaded at no charge.
www.irsst.qc.ca

To obtain the latest information on the research carried out or funded by the IRSST, subscribe to our publications:

- *Prévention au travail,* the free magazine published jointly by the IRSST and the CNESST (preventionautravail.com)
- InfoIRSST, the Institute's electronic newsletter

Legal Deposit
Bibliothèque et Archives nationales du Québec
2020
ISBN : 978-2-89797-070-3
ISSN : 0820-8395

IRSST – Communications and Knowledge Transfer Division
505 De Maisonneuve Blvd. West
Montréal, Québec
H3A 3C2
Phone: 514 288-1551
publications@irsst.qc.ca
www.irsst.qc.ca

WebExpo
Towards a Better Interpretation of Measurements of Occupational Exposure to Chemicals in the Workplace

Jérôme Lavoué[1,2], Lawrence Joseph[3], Tracy L. Kirkham[4], France Labrèche[5], Gautier Mater[6], Frédéric Clerc[6]

[1] Département de santé environnementale et santé au travail, École de santé publique, Université de Montréal

[2] Centre de recherche du CHUM

[3] Division of Clinical Epidemiology, McGill University Health Center

[4] Dalla Lana School of Public Health, University of Toronto

[5] IRSST

[6] Métrologie des polluants, INRS

STUDIES AND RESEARCH PROJECTS

R-1065

This study was funded by the IRSST. The conclusions and recommendations are solely those of the authors.

PEER REVIEW
In compliance with IRSST policy, the research results
published in this document have been peer-reviewed.

ACKNOWLEDGEMENTS

We would like to acknowledge the invaluable contribution of the three programmers involved in this project: Patrick Bélisle, François Lemay, and Daniel Margulius. We would also like to thank the members of the stakeholder and expert committees, for their availability and insightful comments and suggestions.

ABSTRACT

A significant part of industrial hygiene activities is the measurement of workers' occupational exposure levels. Considerable spatial and temporal variability is usually observed in most exposure assessment surveys, frequently with up to 10-fold variations in exposure intensity, despite apparently similar conditions. This has historically represented an important challenge to the interpretation of measured levels with regard to comparison with occupational exposure limits (OELs). There now exists a consensus framework, progressively developed during the last two decades, for the analysis of exposure levels related to exposure limits. Within this framework, exposure levels are assumed to follow, at least approximately, a lognormal distribution. Several parameters from the underlying distribution, deemed associated with health risk, are estimated from a number of measurements and are interpreted relative to the OEL.

These developments, although permitting a better assessment of risk compared to historical approaches, have not been widely adopted by industrial hygiene practitioners, and involve notions of statistics not usually taught in traditional education programs. Moreover they require calculations not usually feasible with common tools such as calculators or spreadsheet programs. While some specific tools have been developed over the years, usually through volunteer initiatives, most are lacking in some areas, be it accessibility, functionality, user-friendliness or complexity. In addition, uncertainty in parameter estimates has mostly been taken into account through formal hypothesis tests or the calculation of confidence intervals, the results of which are not easily conveyed to decision makers, hampering the ability of practitioners to efficiently communicate risk. Finally, available tools are standalone, and are not easily integrated within an existing data management structure.

The WebExpo project aimed at improving current practices in the interpretation of occupational exposure levels through the creation of a library of algorithmic solutions to frequently asked risk assessment questions in industrial hygiene. Most of these questions require the estimation of parameters from one or several distributions. WebExpo has utilized Bayesian statistics to perform these tasks. Bayesian methods were chosen due to two main advantages: first, they provide inferences in direct probabilistic terms (e.g. what are the odds that…?), facilitating risk communication. Second, they tackle methodological issues rarely taken into account, such as the data reported as not detected (a frequent concern). The three specific objectives of WebExpo included: 1) to assess current needs in calculation, documentation and risk communication associated with the interpretation of occupational exposure measurement data, 2) to create a library of computer programming codes based on Bayesian statistics that answers a set of data interpretation questions elaborated in specific objective 1, 3) to create prototype tools using the code from specific objective 2 that computes industrial hygiene statistics and answers to needs established in specific objective 1.

Specific objective 1 was achieved through a review of international guidelines and recent relevant literature, complemented by meetings with stakeholder and expert committees. Specific objective 2 was achieved through creating Bayesian solutions to the list of calculations finalised in step 1, implementing these algorithms in statistical code, and translating the code into programming language. Finally, the programming algorithms were used to create functioning data analysis prototypes able to showcase the calculation and useable as a starting point for the creation of practical data analysis tools.

The list of relevant calculations resulting from specific objective 1, and later implemented mathematically as well as in the form of algorithms and prototypes, included two main avenues. The first involved estimating parameters from one distribution, i.e., the traditional “similar exposure group” approach. The measurements are assumed to come from a distribution of exposures shared by a group of workers performing similar tasks. As an illustration, this model permits to answer the question: “What is the probability that unmeasured exposures for this group exceed the OEL more than 5% of the time?” The second model extends the first model by permitting to estimate to what extent a group of workers does or does not share similar exposures. The global exposure variability is split into within- and between-worker variabilities. It is possible to assess the group risk but also whether some individual workers might experience higher risk than the group. As an illustration, this model permits to answer the question: “Although group exposure seems acceptable, what is the probability that a randomly selected worker might experience exposure exceeding the OEL more than 5% of the time?” All models include the seemless treatment of non-detects and take into account measurement errors associated with the observations.

The resulting algorithms are available in R, aimed at academics, C#, for standalone offline or server-based applications, or JavaScript, for web-based applications. They include data entry, core Bayesian estimation, numerical data interpretation modules, as well as a limited user interface for the C# and JavaScript prototypes. The code is publicly available under the open source licence Apache 2.0 to allow users to build their own applications.

The WebExpo project should result in a comprehensive toolbox, available to the industrial hygiene community for the interpretation of occupational exposure levels, with the added flexibility for users to build or adapt their own software instead of using a new one.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

AIHA: American Industrial Hygiene Association

AM: Arithmetic mean

BOHS: British Occupational Hygiene Society

CEN: Comité européen de normalisation (European Committee for Standardization)

CrI : Credible interval

CV, CV_e: Coefficient of variation

GM: Geometric mean

GSD: Geometric standard variation

IH: Industrial hygiene/industrial hygienist

INRS : Institut National de Recherche et de Sécurité

LIMS: Laboratory information management system

LOQ: Level of quantitation

LTA-OEL: Long-term average OEL

MCMC: Markov Chain Monte Carlo

NIOSH: National Institute of Occupational Safety and Health

NVvA : Dutch Occupational Hygiene Society

OEL: Occupational exposure limit

OSHA: Occupational Safety and Health Administration

P95: 95^{th} percentile

RSPSAT: Réseau de santé publique en santé au travail

SEG: Similar exposure group

TWA: Time-weighted average

1. INTRODUCTION

1.1 Control and management of airborne chemical exposures in the workplace

Industrial hygiene's main goal is to identify hazards, and evaluate, control and manage risks in the workplace. A significant part of these activities relies on acquiring knowledge of exposure levels experienced by workers through breathing air contaminated with chemicals. Such exposure assessment can be required for several purposes. Chemical risk assessment in the workplace often relies on comparing workers' exposures to occupational exposure limits (OELs) or guidelines set by various organizations or governing bodies. Exposure assessment can also be performed to understand the factors that determine exposure intensity in order to target intervention.

While some exposure assessment needs can be met through indirect methods such as control banding or the use of mathematical models, in many situations direct measurement of exposure through sampling and analysis of the air breathed by workers is necessary.

1.2 Implications of environmental variability on exposure assessment: the exposure profile

When measuring exposures in the workplace, the aim is generally not to acquire knowledge only about the particular period sampled, but to infer from this period what is usually happening under the same circumstances. Hence the objective is to obtain a representative picture of exposures corresponding to a set of conditions. For example, when evaluating a worker's exposure level for a full work shift, one would want to use that exposure information to gain knowledge about all other unmeasured days. Indeed, it's the ensemble of exposure-days experienced by the workers, the so-called exposure profile, that reflects risk.

It was recognized early on that exposure levels in the workplace vary considerably across location, time, and workers. Even measurements integrated over a full work shift and repeatedly taken for a similar work situation can often show 10-fold variations from one day to another (N. Esmen, 1979; Kumagai & Matsunaga, 1995; Oldham, 1953; Rappaport, 2000). Therefore, exposure corresponding to a particular situation (e.g. painting metal parts in a body shop) cannot be described by a typical single concentration value. It is rather characterized by an ensemble of different exposure levels due to minute variations in many determining factors (e.g. surface being painted, open/closed doors, air movement, workers' experience, etc.). Estimating such variability is essential in order to draw an accurate portrait of the exposure profile and reliably assess risk.

1.3 Statistics in industrial hygiene: the lognormal distribution

Statistical methods developed to address the challenge posed by environmental variability started to appear in the scientific literature in the 1960s (Breslin, Ong, Glauberman, George, & Leclare, 1967; Kerr, 1962; Roach, 1966), and progressively coalesced into guidelines published by several institutions that inform the practice of occupational hygiene in various countries. Namely, the American Industrial Hygiene Association (AIHA) (Hawkins, Norwood, & Rock, 1991), the National Institute of Occupational Safety and Health (NIOSH) (N. A. Leidel, Busch, & Lynch, 1977), the British and Dutch occupational health societies (BOHS and NVvA respectively) (BOHS-NVvA, 2011; BOHS Technology Committee Working Group, 1993) and the

Institut National de Recherche et de Sécurité (INRS) (INRS, 2018) in France published guidelines for the comparison of exposure levels to exposure limits. The European community recently updated recommendations (CEN, 2018) initially issued in 1995 (CEN, 1995).

These methods assume that environmental variability is adequately modelled by the lognormal distribution model. Under this model, it is assumed, for a given exposure group (e.g. stainless steel welders in a part-manufacturing facility), that the ensemble of exposure levels experienced by workers in this group over a period of relatively stable work conditions (e.g. a year), hereafter referred to as the exposure distribution, follows a lognormal model. Then, when a set of measurements are taken, they are assumed to form a random sample from this exposure distribution. It is therefore possible to draw inferences on the exposure distribution from this sample, including measured and unmeasured days. There is now a large body of evidence suggesting that the lognormal model is a reasonable default assumption for most exposure situations involving vapors and aerosols (N. A. Esmen & Hammad, 1977; Kumagai & Matsunaga, 1995; Oldham, 1953; Roach, 1977).

1.4 Current best practice in industrial hygiene measurement data interpretation

The recommended approaches to comparing measured exposure levels to an exposure limit have significantly evolved during the last 30 years. The initial guideline utilizing a statistical framework for interpretation, proposed by NIOSH in 1977, recommended that exposures should be controlled so that less than 5% of exposure levels experienced by a worker exceed the OEL (N. A. Leidel *et al.*, 1977) (i.e., the 'exceedance fraction' should be <5%; a concept also found in the more recent European standard (CEN, 2018)). At the time, NIOSH proposed to verify this by comparing a single exposure value to an action limit set at half the OEL. Although this proposition was based on statistical grounds and provided a practical way to perform risk assessment, it was later recognised that comparing one measurement to half the OEL did not permit to ensure that 95% of unmeasured exposures would be under the OEL (Buringh & Lanting, 1991; Lyles & Kupper, 1996; Rappaport, 1984; Tornero-Velez, Symanski, Kromhout, Yu, & Rappaport, 1997). Further methodological developments in the following decades identified several risk metrics based on the lognormal distribution (see below) which were embraced in a 2008 workshop about the update of guidelines from NIOSH (Ramachandran, 2008). In all cases, the exposure distribution is the ensemble of exposure concentrations experienced by a group of workers assumed to share similar exposure conditions (i.e., similar exposure group).

1.4.1 Proportion of exposures exceeding the OEL (exceedance fraction)

This metric is directly related to NIOSH's early proposal that less than 5% of exposures should exceed the OEL (N. A. Leidel *et al.*, 1977; N. Leidel, Busch, & Crouse, 1975). Applied to shift-long exposures, the exposure distribution of interest would comprise all time-weighted-averaged (TWA) exposures occurring during a period of stable conditions, typically a year. One would then collect a random sample from this exposure distribution and estimate the exceedance fraction, i.e., the proportion of days expected to be associated with exposure over the OEL. The calculation of exceedance fraction is recommended by the INRS in France, the British and Dutch occupational hygiene societies (BOHS/NVvA) and the European committee for standardization (CEN), and forms the basis of the current French regulation (BOHS-NVvA, 2011; CEN, 2018; INRS, 2018; République française, 2009). Because the estimate of the

exceedance fraction is made from a sample of the exposure distribution, uncertainty has to be taken into account through the calculation of confidence limits around the estimate. In the above recommendation, compliance with an OEL amounts to showing that the 70% upper confidence limit of the exceedance fraction is smaller than 5%. Put in simpler terms, one has to demonstrate with at least 70% certainty that less than 5% of exposures are over the OEL. Comparing the exceedance fraction to 5% is numerically equivalent to comparing the estimated 95th percentile of the underlying distribution to the OEL (Clerc & Vincent, 2014). The latter calculation is recommended in the current guidelines from AIHA (Jahn, Bullock, & Ignacio, 2015), with the associated determination of a 95% upper confidence limit (as opposed to 70% above).

1.4.2 Long term arithmetic mean of the exposure distribution

Toxicokinetic models have shown that the arithmetic mean (AM) of the long term distribution of exposure levels is a more adequate risk metric for evaluating cumulative damage from exposure to most chronic toxicants compared with metrics reflecting the upper tail of the distribution (as the exceedance fraction) (Rappaport, 1991). Within this framework, one would make a number of measurements, estimate the arithmetic mean of the underlying exposure distribution as well as the confidence limits around the point estimate, and compare them with the OEL. There has been some debate about the use of this metric, as it is less conservative than the exceedance metric (i.e., the AM just at the OEL for a typical lognormal distribution would correspond to approximately 30% exceedance) (P Hewett, 1997; Lyles & Kupper, 1996; Tornero-Velez *et al.*, 1997). The current guidelines from the AIHA recommend this approach in cases where the exposure limit has explicitly been defined as a long term cumulative dose index ('LTA-OEL, Long term average OEL') (Jahn *et al.*, 2015).

1.4.3 Probability of individual overexposure

Following seminal work by Kromhout, Rappaport and Symanski (Kromhout, Symanski, & Rappaport, 1993; Rappaport, Kromhout, & Symanski, 1993), it was recognized in the late 1990s that the traditional practice of grouping workers performing similar tasks in the same environment into so-called homogeneous exposure groups could result in an underestimation of the risk for some members of the group. Thus, despite an acceptable group exposure distribution, high variability of exposure between workers could result in a distinct possibility that some workers would have an unacceptable individual exposure distribution. This was notably reflected in the AIHA guidelines, where "homogeneous exposure group" was replaced with "similar exposure group" (SEG) in most recent editions. The AIHA also recommends using analysis of variance methods when enough data are available to assess empirically whether the group is indeed "homogeneous" (Hawkins *et al.*, 1991; Ignacio & Bullock, 2008; Mulhausen & Diamano, 1998). This concept is an integral part of the most recent guidelines by the BOHS-NVvA guideline "Testing Compliance with Occupational Exposure Limits for Airborne Substances" (BOHS-NVvA, 2011). The guideline is a 2-step process. The exposure group distribution is first evaluated to assess whether less than 5% of exposures are above the OEL (similar to the European recommendation mentioned above). If group risk is acceptable, then the guideline requires testing to determine whether there is significant exposure variability between workers within the group to estimate the probability that a random worker's exposure distribution would correspond to an exceedance fraction above 5%. If this probability is estimated greater than 20%, the guideline's diagnosis is "failure to comply". In their early recommendations, Rappaport *et al.* and Lyles *et al.* suggested determining the probability that a random worker would have his own arithmetic mean above the OEL, and comparing it to a

threshold of 10% (Lyles, Kupper, & Rappaport, 1997b, 1997a; Rappaport, Lyles, & Kupper, 1995). Within this framework, the extent of between-worker variation can be measured through calculating the within-worker correlation coefficient rho (rho is close to 1 when between-worker differences are high: the more different workers are, the more measurements for the same worker are close to each other relative to those of other workers), or by the so-called R ratio. The R ratio, initially proposed by Rappaport *et al.*, approximately represents the ratio of the geometric mean (GM) of the most exposed worker to the GM of the least exposed worker (Rappaport *et al.*, 1993).

In summary, current guidelines in industrial hygiene data interpretation recommend four main metrics as the most relevant for risk assessment: exceedance fraction, 95th percentile, arithmetic mean for long-term averaged OELs, and probability of individual overexposure (overexposure defined as an individual's arithmetic mean > OEL or 95th percentile > OEL). These metrics can also be used for analyses other than comparison with OELs, including evaluation of the effect of exposure determinants (e.g., effect of an intervention).

1.5 Bayesian methods to interpret occupational exposure data

Bayesian statistics represent an alternative method for drawing inference from a sample compared to the traditional 'frequentist' approach. In Bayesian inference, one establishes prior beliefs about a set of unknown parameters in the form of probability distributions. Bayes' theorem is then used to update these beliefs with empirical observations, resulting in "posterior" probability distributions for the parameters of interest. While the theory was established during the 18th century, Bayesian methods have only gained popularity relatively recently with the advent of high-computing power. Bayesian statistics have been proposed for use in occupational hygiene because they permit the integration of expert judgment (in the form of prior beliefs) into measurement data (S. Banerjee, Ramachandran, Vadali, & Sahmel, 2014; Paul Hewett, Logan, Mulhausen, Ramachandran, & Banerjee, 2006; Ramachandran & Vincent, 1999; Sottas *et al.*, 2009).

1.5.1 Principle of Bayesian data analysis

Bayesian data analysis begins with stating prior probability distributions for the parameters of a model, which represents current knowledge available about these parameters, prior to considering the current dataset. These prior densities are then updated with information from the current dataset through the likelihood function, leading to the posterior distribution for the estimated parameters (Gelman, 2013; McElreath, 2016). The posterior density represents all available current knowledge, having combined past information with that in the current dataset. Thus all inferences are drawn from this posterior density.

Bayes' theorem is used to perform this update, and can be stated simply as follows:

$$posterior = \frac{prior * likelihood}{normalizing\ constant}$$

The normalizing constant serves only to ensure that the posterior integrates to one.

The prior density represents the information that is available prior to the analysis of the current dataset: it is possible to set these prior densities to have very low information content (these are labelled non-informative or weakly informative) or high information content. In the context of IH, informative priors are attractive because the information added to actual measurements compensates, to a certain extent, for the small sample sizes commonly encountered in industrial hygiene evaluations (Sudipto Banerjee, Ramachandran, Vadali, & Sahmel, 2014; Paul Hewett *et al.*, 2006; Ramachandran & Vincent, 1999; Sottas *et al.*, 2009).

1.5.2 Bayesian data analysis in occupational health

Among the the earliest mentions of Bayesian methods in our field, Ramachandran and Vincent proposed to inform historical exposure reconstruction with expert judgment (Ramachandran & Vincent, 1999). Hewett *et al.* proposed a tool to evaluate the probability of the 95th percentile of the exposure distribution being in each of the AIHA exposure control categories (Paul Hewett *et al.*, 2006). Sottas *et al.* proposed a tool combining measurements with prior information from expert judgment, an existing exposure database, and a mechanistic model (Sottas *et al.*, 2009). More recently, Banerjee *et al.*, as well as McNally *et al.* proposed Bayesian frameworks for comparing exposure data with OELs (Sudipto Banerjee *et al.*, 2014; McNally *et al.*, 2014). Jones and Burstyn, as well as Quick *et al.* proposed specific prior distributions to use when interpreting measurement data with Bayesian statistics, while Huynh *et al.* compared traditional and Bayesian statistics for the treatment of non-detects (Huynh *et al.*, 2016; Jones & Burstyn, 2017; Quick, Huynh, & Ramachandran, 2017). Most recently, Remy-Martin *et al.* and Groth *et al.* proposed Bayesian solutions to handle bivariate censored data for linear regression (Groth *et al.*, 2017; Martin Remy & Wild, 2017).

The first Bayesian tool for data analysis in IH involved defining prior information in the form of prior probabilities for the 95th percentile of the exposure distribution to be in each of the AIHA risk management categories, where probabilities were updated with the observed data through Bayesian analysis (Paul Hewett *et al.*, 2006). Some years later, a bayesian model was described for the Advanced REACH tool (ART[1]), where a mechanistic model combined with a database of measurements associated with various exposure scenarios informs the prior distributions (McNally *et al.*, 2014). Since then, several other options have been proposed to create informative priors, although none to our knowledge have been implemented in practical tools. They include using mechanistic models (Zhang, Banerjee, Yang, Lungu, & Ramachandran, 2009), existing relevant studies (Quick, Huynh, & Ramachandran, 2017), simple software such as EXCEL to estimate posterior distributions (Jones & Burstyn, 2017), and priors from historical exposure databases (Sottas *et al.*, 2009). The traditional Bayesian approach recommends assessing robustness across a range of different priors to widen the interpretation of an analysis (Gelman, 2013) as it will apply to a wider variety of interpretations of prior data. For realistic sample sizes in our field (5-10 observations), informative priors such as those described above will typically have a non-trivial effect on the final exposure estimates (Jones & Burstyn, 2017).

There are other significant advantages to using Bayesian statistics to interpret industrial hygiene data. Bayesian inference is probabilistic in nature, therefore instead of a hypothesis test or a confidence interval, whose correct interpretations are sometimes difficult to convey to the layman, Bayesian analysis provides answers to questions in the direct form of “what is the probability that” (e.g., what is the probability that this group is overexposed more than 5% of

[1] https://www.advancedreachtool.com/

days; or, what is the probability that this intervention reduced exposure levels by at least 50%). This facilitates risk communication of complex concepts to management and workers. Furthermore, two technical challenges currently not appropriately tackled by traditional approaches, namely the handling of non-detects and incorporating measurement error into an assessment, are easily integrated into a Bayesian approach (Espino-Hernandez, Gustafson, & Burstyn, 2011; Mcbride, Williams, & Creason, 2007; McNally *et al.*, 2014; Morton, Cotton, Cocker, & Warren, 2010; Pilote *et al.*, 2000; Wild, Hordan, Leplay, & Vincent, 1996).

The Bayesian framework therefore appears to be a very promising avenue to improve data analysis and interpretation in industrial hygiene. Unfortunately, its implementation is currently out of reach for most practitioners, as running Bayesian computations requires advanced software and technical knowledge, usually limited to academic specialists.

1.6 Treatment of non-detects in occupational health measurement data interpretation

In 1990, Hornung and Reed wrote that the reduction in exposure levels since the 70s, only partially mitigated by gradually improving analytical methods, increased the proportion of exposure data reported as non-detected (Hornung & Reed, 1990). More recently, Lavoué *et al.* reported 60% of non-detects in 1.4M measurements in the database from the Salt Lake City laboratory of the Occupational Safety and Health Administration (OSHA) in the US, which includes a majority of the samples taken by OSHA officers since 1979 (Lavoue, Friesen, & Burstyn, 2013). Sarazin *et al.* reported 40% of non-detects among 0.5M measurements in the IRSST laboratory information management system (LIMS) database, which contains analysis results from samples collected since 1985 in Quebec by governmental industrial hygienists (Sarazin, Labrèche, Lesage, & Lavoué, 2018).

There is ample evidence that eliminating non-detects, or replacing them with any fixed value, will bias the estimation of most parameters of interest (D. R. Helsel, 2012; D Helsel, 2005). The amount of error increases with increasing proportion of non-detects, and is particularly severe when conducting statistical tests or calculating confidence intervals. Despite the pervasiveness of non-detects and the potential impact on data interpretation, few methodologies have been proposed in our field, even with editorials appearing in the Annals of Occupational Hygiene pleading for advances (Dennis Helsel, 2010; T. L. Ogden, 2010). Significant progress has been reported recently (Flynn, 2010; Ganser & Hewett, 2010; Groth *et al.*, 2017; Krishnamoorthy, Mallick, & Mathew, 2009; Martin Remy & Wild, 2017), with several simulation studies comparing approaches (Paul Hewett & Ganser, 2007; Huynh *et al.*, 2014, 2016).

Bayesian methods are optimally suited for this challenge, since they allow for multiple censoring points, and they accurately estimate the inherent uncertainty when data values are known only up to an interval (Huynh *et al.*, 2016). These recent developments have unfortunately not yet been implemented in practical data analysis tools.

1.7 Measurement error in occupational health measurement data interpretation

In addition to the variability in exposure levels themselves, each value in a set of exposure measurements is associated with an error due to sampling and analysis. This error, usually expressed as a coefficient of variation (CV) is taken into account when interpreting a single

exposure measurement, traditionally to assess whether the actual underlying exposure concentration was above or below an OEL. As an illustration, Leidel and Bush provide various formulas for estimating confidence intervals for a single time-weighted average value based on sampling and analytical error (N. A. Leidel & Busch, 2000). However, measurement error has not been considered when interpreting a set of exposure measurements to estimate parameters of the distribution of exposure levels. The challenge for this kind of analysis is that while environmental variability is usually modeled by a lognormal probability distribution, measurement error corresponding to sampling and analysis is rather modelled by a normal distribution (Ashley & Bartley, 2004; Bartley, 2001; Bartley & Lidén, 2008). This renders a combined analysis intractable with traditional statistics.

The current practice of not considering measurement error when interpreting IH datasets is supported by two studies that approximated the normal probability distribution for measurement error with a lognormal probability distribution (Grzebyk & Sandino, 2005; Nicas, Simmons, & Spear, 1991). Nicas *et al.* (Nicas *et al.*, 1991) estimated the contribution of measurement error to the total observed variability, while Grzebyk and Sandino (Grzebyk & Sandino, 2005) derived equations for the bias caused to the estimation of geometric mean (GM) and geometric standard deviation (GSD). Both concluded that measurement error had a negligible contribution when the corresponding CV is <30% and environmental variability is high (GSD>2). However, variability in some workplaces can be low, and some sampling methods have a considerable measurement error. Moreover, no approach yet has been proposed to estimate the error on a full-shift value calculated from a sequence of partial samples summing to less than the whole shift, which might be higher than the typical sampling and analysis error. Finally, neither Grzebyk and Sandino nor Nicas *et al.* estimated the impact on the decision metrics described above. Hence, while bias in the estimation of GSD (not taking measurement error into account would typically cause an overestimation of the true GSD) might appear small, the actual impact on the upper confidence limit for the 95th percentile of the exposure distribution (often used for decision making) might be significant.

As in the case of non-detects, Bayesian statistics represent a promising alternative to frequentist statistics, as they can account for measurement error in a flexible manner (Espino-Hernandez *et al.*, 2011; Morton *et al.*, 2010; Pilote *et al.*, 2000).

1.8 Challenges with data interpretation and risk communication

Recent studies on expert judgment have shown that industrial hygienists performed better at gauging exposure levels when taught specific courses about lognormal statistics (P. Logan, Ramachandran, Mulhausen, & Hewett, 2009; P. W. Logan, Ramachandran, Mulhausen, Banerjee, & Hewett, 2011). In Quebec, modern approaches to data interpretation were reviewed and summarized in a recent IRSST report (Drolet *et al.*, 2013). The authors specifically pointed out that these approaches require statistical notions and calculation tools not widespread in the field.

Risk communication is also a challenge that would benefit from any improvement as statistical concepts often appear obscure to decision makers and workers. For instance, it is possible to have an exposure situation where a set of measurements are all under the OEL, but the estimated proportion of exposures expected to be over the OEL during unmeasured days would be much greater than the generally accepted 5%. This particular assessment would probably seem counter-intuitive to an uninformed audience but seem intuitively reasonable once one realizes that having several observations just below the cutoff value in a density with a long tail

(such as the lognormal) can result in a substantial probability of being in that tail, and hence above the OEL. The difficulty and lack of tools to efficiently communicate statistical results in a convincing way to non-specialists may also explain the slow appropriation of modern guidelines by practitioners in the field.

1.9 Numerical and statistical analysis needs for the interpretation of occupational exposure data

Statistical procedures for lognormal parameters and their uncertainty are not described in standard statistical textbooks, which are mostly centered on the normal distribution. Hence, they have been gradually developed since the 1960s onward and are evolving. While these developments trickled down from research papers into guidelines from industrial hygiene associations over time, their implementation can be complicated, making it difficult for practitioners who may lack the statistical knowledge and tools to perform such calculations. In Quebec, the Sampling guide for air contaminants in the workplace (Drolet & Beauchamp, 2013) is referred to by the regulation as a reference on the level of accuracy required for how to assess exposure regulatory compliance to OELs. The guide provides detailed instructions on how to compare one measurement to the OEL in order to determine whether exposure on the measured day was compliant, which is essential for regulatory compliance officers. However, it does not include comprehensive documentation of the lognormal distribution and the associated risk metrics. We identified only five available practical evaluation tools that focus on the estimation of the statistics necessary for risk assessment of airborne chemicals in industrial hygiene (i.e., industrial hygiene statistics): IHSTAT[2] (free Excel worksheet), Altrex Chimie[3] (free standalone downloadable software), IHData analyst[4] (for free software), BW_Stat[5] (free Excel worksheet) and HYGINIST[6] (free standalone downloadable software). We should also mention ProUCL[7], made available by the US Environmental Protection Agency, which is a generic toolset for environmental contamination and can be applied to occupational exposure datasets. Additionaly, the ART tool mentioned in section 1.5.2, while focusing more specifically on the risk assessment framework defined by the REACH regulation, allows estimating percentiles of an exposure distribution and associated uncertainty (McNally *et al.*, 2014). The IH specific tools mentioned above share similarities, and all allow assessment of risk of overexposure based on at least one or several metrics. As such they represent an important step towards making industrial hygiene statistics more accessible. However, none provides a single integrated and comprehensive solution to lognormal data interpretation. Most notable limitations include lack of adequate treatment of non-detects, data interpretation outside of assessment of overexposure (e.g., effect of an intervention), and support for probabilistic risk communication.

In addition, many institutions and private companies performing exposure measurement routinely maintain their own exposure databank; however, none of the tools presented above can be easily integrated in an existing data management system. Therefore, in order to be able to perform the relevant calculations, it is necessary to export data into the existing tools and

[2] https://www.aiha.org/get-involved/VolunteerGroups/Pages/Exposure-Assessment-Strategies-Committee.aspx
[3] http://www.inrs.fr/accueil/produits/mediatheque/doc/outils.html?refINRS=outil13
[4] https://www.easinc.co/ihda-software/
[5] https://www.bsoh.be/?q=en/node/89
[6] http://www.tsac.nl/hyginist.html
[7] https://www.epa.gov/land-research/proucl-software

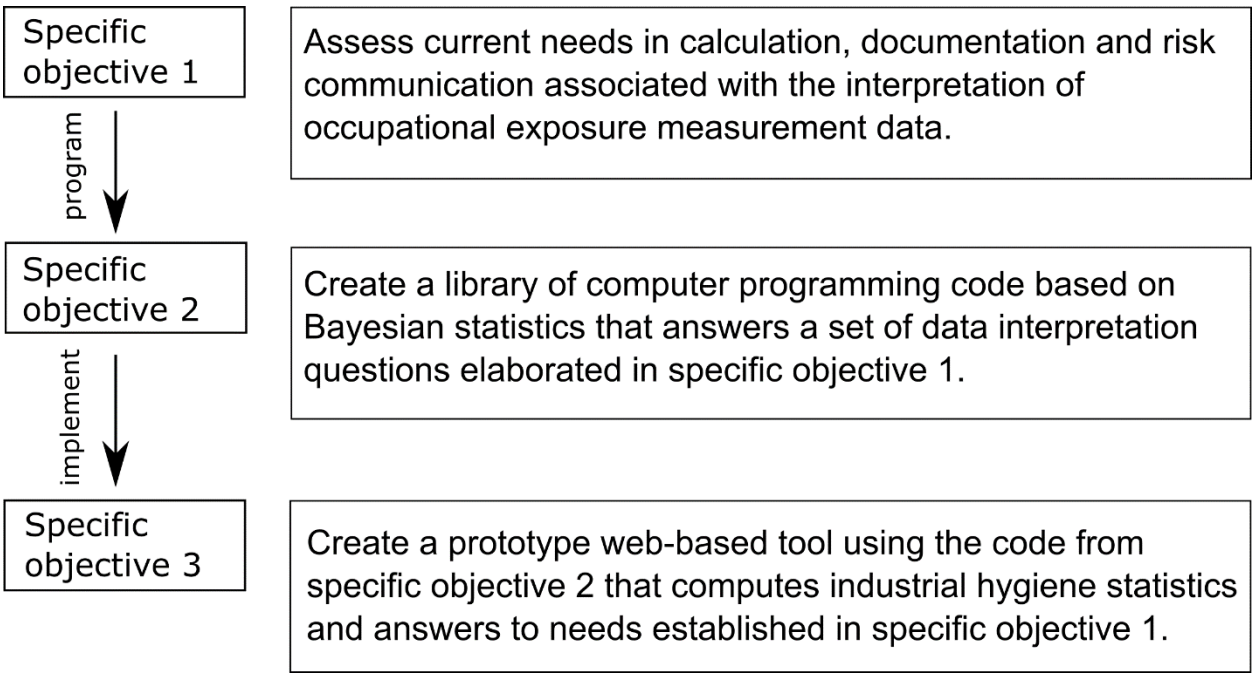
reimport the results. This signifies a need to provide a hygiene statistics toolbox that would facilitate the process of programming necessary calculations into existing systems.

1.10 Summary of knowledge gaps and needs

Considerable spatial and temporal variability observed in levels of exposure has historically represented an important challenge to their interpretation. A consensus framework now exists for their analysis based on the lognormal distribution. These developments, although permitting a better assessment of risk compared to historical approaches, have not been widely adopted by industrial hygiene practitioners. Indeed, they involve statistical notions not usually taught in traditional training programs and require calculations not usually feasible with simple common tools (e.g., calculators or spreadsheets). The few specific tools currently available are an important step forward but they do not yet present a comprehensive answer to practitioners' needs. Moreover, available tools are standalone, and are not easily amenable to integration within an existing data management structure. Finally, Bayesian methods represent a very promising approach to data interpretation in industrial hygiene but are currently not accessible to most practitioners. In conclusion, to support the adoption in the field of modern guidelines for industrial hygiene data interpretation and to improve chemical risk assessment practice there is a significant need for better knowledge translation, and for accessible and comprehensive data analysis tools.

2. RESEARCH OBJECTIVES

The WebExpo project aimed at improving transfer of current best practices in the interpretation of occupational exposure levels for risk assessment into the field of occupational hygiene.



3. METHODS

3.1 Specific objective 1: Establishing current needs in calculations, documentation and risk communication

Based on the review presented in section 1 it is possible to produce a tentative list of core features that should be minimally present in a comprehensive data interpretation tool:

Metrics relevant to group assessment of overexposure:

- Exceedance fraction
- 95^{th} percentile
- Arithmetic mean

Metrics relevant to individual assessment of overexposure:

- Between- and within-worker components of variability
- Within-worker correlation
- R ratio
- Probability that a worker's individual exceedance fraction is too high
- Probability that a worker's individual 95^{th} percentile is too high
- Probability that a worker's individual arithmetic mean is too high

Uncertainty management:

- Calculation of confidence intervals for all metrics above.
- Treatment of non-detects
- Treatment of measurement error

We validated and sought to possibly extend this list through obtaining feedback from stakeholders in the form of two committees.

First, we formed a stakeholder committee made of industrial hygiene practitioners from Quebec. The committee included nine stakeholders: an industrial hygienist (IH) from a consulting company, two IHs from private companies, an IH from IRSST, an IH from the Public health network in occupational health (RSPSAT, Réseau de santé publique en santé au travail), an industrial hygiene technician and an occupational physician from the RSPSAT. The stakeholder committee convened twice for a half-day at the beginning and at the end of the project. The main goal of this committee was to provide feedback and suggestions from the standpoints of Quebec practitioners about needs and obstacles to the use of current data interpretation guidelines.

Second, we created an expert committee with Canadian and international experts in the field of industrial hygiene statistics, with academic and private affiliations (Table A1 in Appendix A). The expert committee convened once for two days at the beginning of the project. The main goal of this committee was to provide feedback and suggestions on methodological choices for the calculation and features that would end up being included in the algorithms.

Both committees helped define the final list of functionalities and calculations covered in WebExpo.

3.2 Specific objective 2: Creation of a library of computer code

This task can be separated into two components. The first one involved setting up theoretical Bayesian solutions to the list of estimation problems finalized in 3.1. As will be detailed in section 3.2.1, this component mainly consisted in transposing to the field of IH statistical techniques that are already available in other domains. The second component involved implementing the calculations in the form of algorithms to ultimately facilitate their use by a large audience, outside of users of specialized statistical packages. As described in section 3.2.2, this required first implementing solutions in the R statistical package, and then translating the R code into languages used for programming web or standalone applications.

3.2.1 Setting Bayesian models for the estimation problems defined in 3.1

The Bayesian models for the WebExpo project were set up based on already published IH literature, available techniques already described in other fields, as well as the expertise of the McGill biostatistics team members (Lawrence Joseph and Patrick Bélisle), who were asked to detail their underlying mathematical basis. We used the models presented in Banerjee *et al.* (S. Banerjee *et al.*, 2014) and McNally *et al.* (McNally *et al.*, 2014), respectively, as a starting point for the group and variance component models, which we extended to include censoring, measurement error, and multiple prior types.

The models in the WebExpo project were anticipated to be too complex and for which the posterior distributions could not be easily written analytically in closed form, i.e., there would be no deterministic equation allowing the estimation. As a consequence, Markov Chain Monte Carlo (MCMC) simulation would be necessary to obtain samples from the posterior distributions from which inferences are made. As an example, let's consider estimating the mean μ of a normal distribution with standard deviation σ. After setting up priors for the unknown parameters and collecting a sample of observations, a typical MCMC output would include, for example, 10 000 random values from the posterior distribution for μ. The point estimate for μ would be the median of these 10 000 values, and the 2.5^{th} and 97.5^{th} percentiles of the 10 000 values would constitute a 95% equal-tailed credible interval (95% CrI). Bayesian credible intervals are interpreted as direct probabilistic statements: the probability that μ is in the interval is 95%, given the current and past information, as represented by the prior and likelihood function used.

One feature of Bayesian data analysis especially useful in our project is the fact that once posterior samples for the model parameters are available, posterior samples for any function derived from them are also immediately available. In the example above, let's assume we obtained 10 000 values for μ and 10 000 values for σ from their joint posterior distribution. It is then straightforward to obtain 10 000 values for the coefficient of variation of the distribution, simply by calculating $CV=\sigma/\mu$ from each pair of samples of μ and σ. The output of the Bayesian models being estimates of basic distributional parameters (e.g., geometric mean and standard deviation), we also developed equations to transform the MCMC chains for these estimates into the relevant metrics selected in 3.1 (e.g., 95th percentile). This was mostly based on existing guidelines and industrial hygiene publications (see section 1 for references). Figure 1 Illustrates the Bayesian estimation process as implemented in WebExpo.

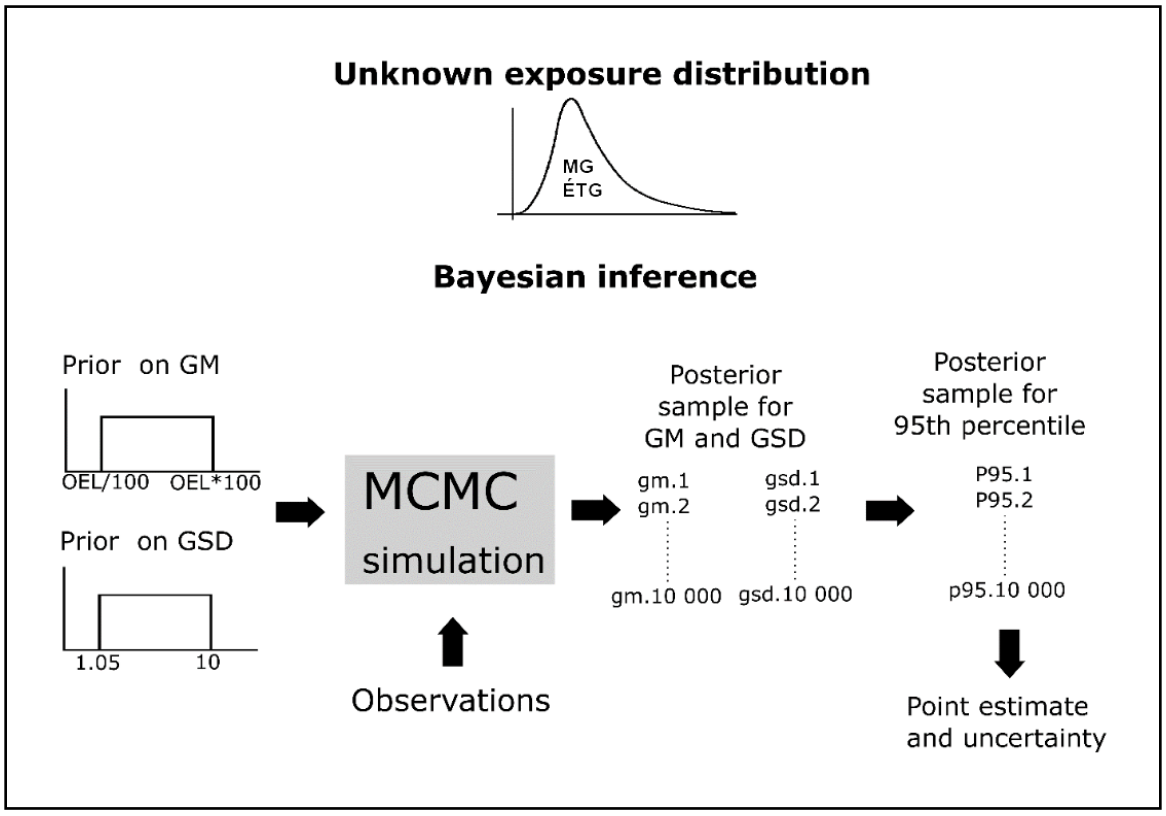


Figure 1. Illustration of the Bayesian analysis framework in WebExpo.

3.2.2 Creation of the library of programming codes

3.2.2.1 General approach

Markov Chain Monte Carlo methods are computer intensive and Bayesian calculations are usually performed with specialized softwares such as Openbugs[8], Winbugs[9], JAGS[10], or STAN[11]. The code for these applications is usually reported in research papers such as Banerjee *et al.,* (2014). However, these programs are too complex for day-to-day use by IH practitioners. Hence implementing the WebExpo algorithm using the above-mentioned software would not ultimately facilitate the practitioners' calculation needs.

We therefore opted for first implementing the WebExpo Bayesian models using a tailor-made MCMC engine written in the free R statistical language (R Core Team, 2014) using basic calculus functions, that would be amenable to subsequent translation into programming languages traditionally used to create practical tools and free of any licensing issue.

The first step in implementing the theoretical algorithms involved creating a library of R code, including initial data formatting, Bayesian calculation, and transformation of the Bayesian

[8] http://www.openbugs.net/w/FrontPage
[9] https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/
[10] http://mcmc-jags.sourceforge.net/
[11] http://mc-stan.org/

functions output into relevant exposure metrics. This would create a library of R code to perform all the calculations in the WebExpo project. However, running this code would still require some expertise in R, as well as a local version of R on the user's computer.

The next step was therefore to translate this library into two computer programming languages, which facilitates running the corresponding routines in other computer environments without having to use statistical analysis packages, and which can be used by programmers to create applications. The first translation was into the web-programming language JavaScript[12]. This translation allows performing Bayesian calculations directly within a standard web environment (i.e., within one's own web browser). The second translation was into the C# (C sharp) language[13]. C# is a prominent programming language used to create standalone applications. It allows the creation of downloadable software capable of performing all the routines created in the project, as well as their integration into existing data management applications.

Finally, the calculations were also coded in R using calls to the JAGS application, a third-party Bayesian engine allowing rapid MCMC simulation for a wide array of models, through the RJAGS package[14]. This set of R+JAGS functions was created to allow R users to perform the calculations described in this report with optimal efficiency, as the pure R models are noticeably slower than their R+JAGS counterpart, especially in the case of measurement error.

3.2.2.2 Quality control

The theory behind the Bayesian models in the WebExpo project is based on published literature, with relatively simple models (single distribution estimation or variance components model). Therefore, we did not attempt to perform simulations, e.g., to verify that our procedure yielded accurate estimates of the geometric mean, or that the credible intervals' coverage was exact. Similarly, the performance of Bayesian treatment of censored data has already been evaluated by others. The theoretical models and R functions were written by two experienced Bayesian statisticians, Lawrence Joseph and Patrick Bélisle.

Our main concern for this project was that the implementation of MCMC algorithms, which rely on random number generation, should provide similar results across different calculation platforms (R, R+JAGS, C#, JavaScript). Even within a single platform, the randomness associated with MCMC implies that running the same analysis multiple times will yield slightly variable results. When the pure R code was translated into C# and JavaScript, differences between platforms could appear because basic mathematical functions are written differently across platforms, or because the random generation mechanisms or rounding procedures are different. For differences between R and R+RJAGS, although the pure R and JAGS scripts start from the same theoretical model, the actual MCMC algorithms are different.

During the translation from pure R to C# and JavaScript, we aimed to ensure that the differences between platforms were minimal, and regular communication took place between the R and C# or JavaScript coders, especially whenever notable differences were observed across platforms. In practice, standard samples (of varying size, distribution, degree of censorship, variability) were analysed using all platforms and the results were compared.

[12] https://en.wikipedia.org/wiki/JavaScript
[13] https://en.wikipedia.org/wiki/C_Sharp_(programming_language)
[14] https://cran.r-project.org/web/packages/rjags/index.html

It was possible to use the same random number generator in C# and in R, which allowed the comparison of the MCMC chains for each standard sample, at each iteration between R and C#, and descriptive statistics of the differences across iterations were used to measure agreement.

For JavaScript, while the same procedure would have been possible in principle, we rather compared quantiles of the posterior samples, namely the 1st, 2.5th, 5th, 25th, 50th, 75th, 95th 97.5th and 99th percentiles of the MCMC chains for all unknown parameters. This procedure was simpler to implement and less time consuming than the iteration-by-iteration approach.

The procedure was simplified to compare R with R+JAGS, as there was less concern about inter-platform differences (both calculations are implemented within R). For each estimation procedure, one standard sample was submitted 50 times to the R and R+RJAGS functions, and we computed the ranges across these repetitions of the point estimates and credible limits for the unknown parameters. The ranges were then compared between R and RJAGS to make sure they were comparable, given the variability observed within each approach.

3.3 Specific objective 3: Creation of prototype tools

The JavaScript and C# libraries described in stage 3.2 were used to create prototype data interpretation tools in both langages. These prototypes, also open source, aim at showcasing the calculations allowed by the algorithms, with a minimal user interface but showing essential numerical results. They both contain a data entry interface where values must be entered for all parameters necessary to the functions alongside the dataset to be analysed. Outputs include the MCMC chains themselves, as well as exposure metric point estimates and credible intervals. No graphical illustration or interpretation of the results is provided. The two prototypes will serve as starting points for the future creation of an IRSST-specific fully-fledged practical data interpretation tool.

4. RESULTS

4.1 Specific objective 1: Establishing current needs in calculations, documentation and risk communication

The literature review presented in section 1 identified two main avenues, which correspond to two different statistical models, for IH data interpretation. The first, hereafter called “SEG analysis”, corresponds to situations when a set of measurements are available for a group of similarly exposed workers (i.e., workers assumed to share the same exposure distribution), or for a single worker. In that case, the analysis involves estimating parameters for a single exposure distribution, from which the exposure metrics such as exceedance fraction can be derived. The second type of analysis, hereafter labelled “between-worker differences”, can be performed when repeated measurements are available for some workers in a group. It allows separating total variability into between- and within-worker components and evaluates the homogeneity of exposure in the group and whether some individual workers might be at risk despite acceptable group exposure. As this dichotomy is reflected throughout the results section, we felt it necessary to introduce it here.

4.1.1 Feedback from the Quebec practitioners committee

A majority of the comments by this committee pertained to recommendations around the design of practical IH calculation tools rather than the actual numerical estimation procedures or relevant exposure metrics. These remarks will be very useful in the next phase of this project for the creation of a tool from the prototypes created in this project. However, they are not relevant for the selection and the algorithmic implementation of calculation routines, and so we did not include the full meeting notes here. We will, however, mention that the practitioners’ committee, in agreement with the experts’ committee (see below), underlined the importance of facilitating risk communication, either through creating easy to understand numerical results, through graphical tools, or through comprehensive documentation accessible to non-specialists.

4.1.2 Feedback from the international expert committee

The final notes of the 2-day meeting held by the expert committee are available in Appendix A. During the meeting, after a general introduction, attendees were presented with the proposed core list of metrics described in 3.1 for discussion. Additional points scheduled for discussion included treatment of non-detects, measurement error, the use of informed Bayesian priors in the calculations, and risk communication. Attendees were free to add any other topic judged relevant.

For both the SEG and between-worker difference analyses, the committee confirmed interest in estimating all metrics in the initial proposal.

Treatment of censored data was deemed essential but could be restricted to left-censored data (as right and interval censorship occur more rarely in IH measurement data).

Including some form of measurement error in the calculations was deemed of interest, but rather as an optional feature, as it was felt that most situations would correspond to a negligible impact.

The committee expressed little interest in formal hypothesis tests to evaluate the adequacy of the lognormal distribution. Hence it appears that below 30-50 data, which would be the majority of situations in IH data analysis, hypothesis tests, or even graphical assessment such as the Q-Q plot, do not provide useful information on distributional shape.

Finally, the committee underlined the importance of creating numerical outputs which would be accessible to non-specialists, and, in particular, which would adequately convey the uncertainty associated with the analyses.

4.1.3 Creation of a probabilistic data interpretation framework

Both committees' statements on the importance of facilitating communication of uncertainty led us to set up an alternative to only using confidence intervals. This framework, described below, basically relies on providing an answer to the question: "What is the probability that this situation corresponds to overexposure?"

Appraisal of uncertainty has traditionally been tackled with confidence intervals and hypothesis tests. For example, to answer the question "Even if the point estimate of the 95th percentile (P95) for a group of workers is < OEL, how sure can we be that the true value is indeed < OEL?" A typical statistical test for this question would state a null hypothesis such as "the true 95th percentile is above the OEL". One would then perform the test and hope to reject the null hypothesis with a small type I error. Alternatively, one would calculate an upper confidence limit and hope that it is lower than the OEL. A common feature of these procedures is that their outcome, based on a pre-selected degree of confidence, is binary. For example, in the case of calculating a 90% upper confidence limit on the 95th percentile, either this limit is lower than the OEL, and we can then be 90% sure that the true value is < OEL, or it is higher, and the conclusion is : "we can't demonstrate with 90% confidence that the true 95th percentile is < OEL".

An alternative and more direct statement about uncertainty could be made. For example, calculating the probability that the true 95th percentile is below the OEL, which should be high (>90% in the above example), or, conversely, the probability that the true 95th percentile is above the OEL, which should be low (<10% in the example above). These statements are both informative, and easy to convey to workers or employers as they provide direct answers to the question "What are the chances that exposure is too high?"

Bayesian analysis naturally permits such direct statements about the degree of uncertainty in the conclusions that can be drawn from data. The probabilistic framework implemented in the WebExpo project involves two stages leading to an estimate of the probability that exposure is not adequately controlled, which we call probability of overexposure, or overexposure risk.

Stage 1 – Definition of overexposure: which characteristic of the exposure distribution corresponds to an unacceptable situation?

As an example, for the SEG analysis, the review presented in section 1 suggested three different definitions of overexposure:

Exceedance fraction ≥ 5%
95th percentile ≥ OEL
Arithmetic mean ≥ OEL

Stage 2 – Analysis of the observed data using the Bayesian models.

In addition to parameter point estimates with credible intervals, the probability that the overexposure criterion is met is estimated from the posterior distribution samples, e.g., the probability that true 95th percentile is ≥ OEL given the data. This quantity, overexposure risk, can be used as a direct input for exposure management: is overexposure risk low enough that it is possible to consider exposure well controlled, or is it high enough that some action should be undertaken (e.g., consider implementing exposure controls)?

Although overexposure risk provides complete information about uncertainty, risk managers often prefer to receive results in the form of a recommendation: does this situation require an intervention, or not? Providing such a recommendation requires setting a threshold for overexposure risk: the situation can then be declared either adequately controlled if overexposure risk is lower than the selected value, or poorly controlled otherwise. That value is called the overexposure risk threshold. The widely accepted value for this threshold is 5%, where the overexposure risk should be lower than 5% to declare a situation acceptable (Jahn *et al.*, 2015).

To illustrate the correspondence between this and more traditional statements of uncertainty, we shall use the example of P95 ≥ OEL as the overexposure criterion. An overexposure risk below 5% is equivalent to “the chances that the true 95th percentile is above the OEL are below 5%”. This means that we are at least 95% certain that the 95th percentile is below the OEL. Finally, this is equivalent to “the 95% upper confidence limit for the 95th percentile is below the OEL” (a more traditional statement). The current British-Dutch and European guidelines, as well as the French regulation, recommend comparing the 70% upper confidence limit for the exceedance fraction to 5%, which is equivalent to a 30% overexposure risk threshold with the overexposure criterion : “exceedance fraction ≥5%” (BOHS-NVvA, 2011; CEN, 2018; T. Ogden & Lavoué, 2012; République Française, 2009). Although the WebExpo algorithms do not perform the actual comparison of overexposure risk with a chosen threshold, such a comparison is trivial to perform using the provided overexposure risk value. Figure 2 below illustrates the correspondence between the traditional use of confidence/credible limits and overexposure risk.

The uncertainty analysis involving calculation of between- and within-worker variabilities has an added layer of complexity as such analyses yield estimates of the probability of individual overexposure, i.e., the probability that a random worker would have an unacceptable individual exposure distribution. A threshold has been proposed, i.e., probability of individual overexposure should be <20% (BOHS-NVvA, 2011). However, as this probability is estimated, it is uncertain. Therefore, instead of only comparing the point estimate of the probability of individual overexposure to 20%, one can evaluate the chances that the true value is ≥20%, i.e.,

the chances that an intervention would be required. A typical example of output would be “Given the data, we estimate that the probability of individual overexposure (overexposure defined as P95 > OEL) is 12% (90% CrI 6-50); the chances that the true value is above the threshold of 20% are 25%.

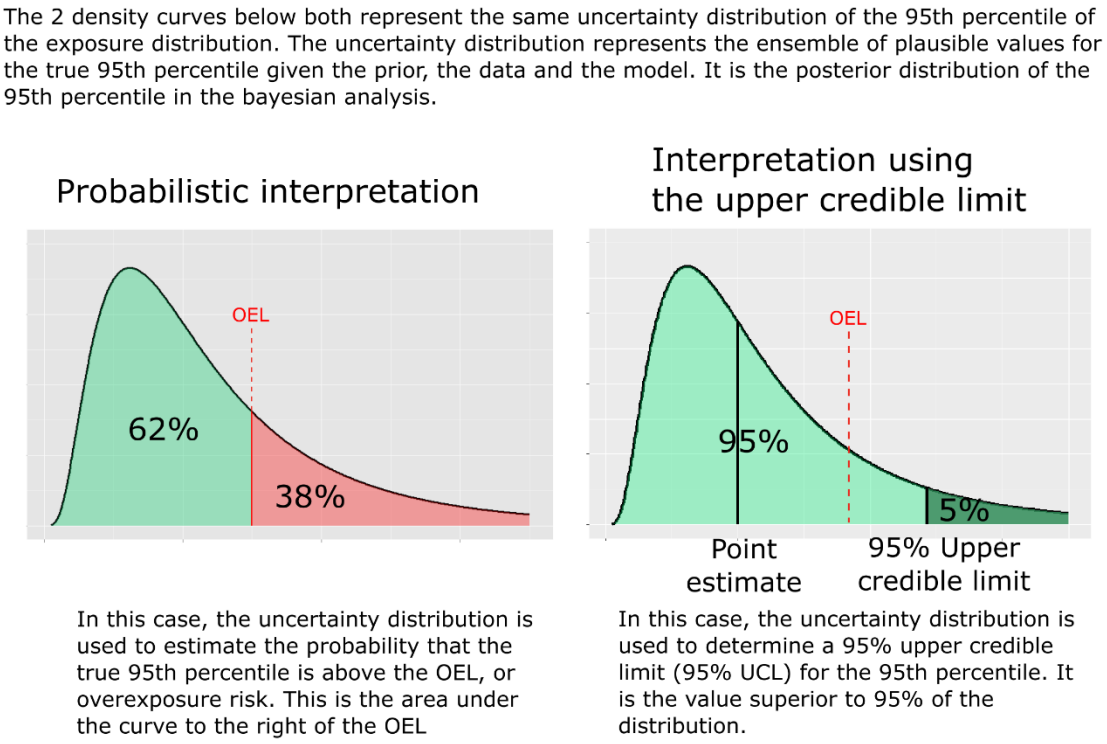


Figure 2. Illustration of the correspondence between overexposure risk and credible limits for the 95th percentile.

4.1.4 Final list of calculations in the WebExpo project

Table 1 presents a glossary of terms and metrics used in the WebExpo project. Table 2 presents the list of metrics that were selected for inclusion, based on the initial review and feedback from the committees.

The reader will note that Table 2 does not include probability of individual overexposure expressed using overexposure defined by exceedance fraction > exceedance threshold. This is because this quantity is equal to Pind.percentile, given the selected percentile corresponds to the exceedance threshold (in default values: 95th percentile/5% exceedance threshold).

Table 1. Glossary of terms

Exceedance fraction	Proportion of exposure levels in the population of interest that are above the exposure limit. Equivalently, probability for a single random exposure value to be above the OEL.
95th percentile	The 95th percentile of a distribution is defined as the value below which lies 95% of the distribution.
Overexposure	Characteristic of an exposure distribution that is unacceptable, i.e., which would trigger preventive action.
Exceedance threshold	Proportion of exposure levels over the OEL used as a threshold to define overexposure (traditionally 5%).
Critical percentile	Percentile of the exposure distribution that will be compared to the OEL to evaluate overexposure (traditionally 95th percentile).
Overexposure risk	Probability that the criteria used to define overexposure is met (e.g., 95th percentile ≥ OEL). Practically: probability of an unacceptable exposure situation.
Overexposure risk threshold	Maximum allowable overexposure risk. This value, chosen *a priori* by the user, is used to create a dichotomy between “adequately controlled” and “poorly controlled” based on the overexposure risk. A traditional value used in the field of statistics would be 5%. The European guideline OEL compliance definition is equivalent to an overexposure risk threshold of 30%.
Probability of individual overexposure	Probability that a random worker within a group would have their individual exposure distribution corresponding to overexposure (e.g., probability that a random worker within a group has his individual 95th percentile above the OEL). Can also be stated as: Proportion of workers with their individual exposure distribution corresponding to overexposure.
Credible interval	While not formally equivalent, Bayesian credible intervals are usually interpreted in a similar way as the more traditional confidence intervals.
R ratio	R ratio has been defined by Rappaport *et al.* as the ratio of the 97.5th percentile of the distribution of workers' individual arithmetic mean divided by the 2.5th percentile of the same distribution.
R difference	Defined by adapting the R ratio to the normal distribution. Difference between the 97.5th percentile of the distribution of workers' individual arithmetic means minus the 2.5th percentile of the same distribution, expressed as a percentage of the group arithmetic mean.

Table 2. Exposure metrics calculated for the lognormal distribution in the WebExpo project

SEG analysis
Distributional parameter estimates (point estimate and credible intervals)
Geometric mean
Geometric standard deviation
Exceedance fraction of the OEL
Percentile of the exposure distribution (i.e., critical percentile, default 95%)
Arithmetic mean of the exposure distribution
Decision on Exposure Acceptability (overexposure risk)
Probability that exceedance fraction ≥ exceedance threshold (default 5%)
Probability that critical percentile (default 95%) ≥ OEL
Probability that arithmetic mean ≥ OEL
Between-worker differences*
Distributional parameter estimates (point estimate and credible intervals)
Group geometric mean
Within-worker geometric standard deviation
Between-worker geometric standard deviation
Within-worker correlation coefficient (rho)
Probability that rho is ≥ threshold (Prob.rho.overX)
R ratio (R.ratio)
Probability that R is ≥ 2 (Prob.R.over2, threshold to define heterogenous groups in Kromhout *et al.*, 1993)
Probability that R is ≥ 10 (Prob.R.over10, threshold to define very heterogenous groups in Kromhout *et al.*, 1993)
Parameters quantifying the possibility that some workers are overexposed (probability of individual overexposure)
Proportion of workers with their individual critical percentile ≥ OEL (Prob.ind.overexpo.perc)
Proportion of workers with their individual arithmetic mean ≥ OEL (Prob.ind.overexpo.am)
Probability that the true value for Prob.ind.overexpo.perc is above a threshold (Prob.ind.overexpo.perc.overX, default 20%)
Probability that the true value for Prob.ind.overexpo.am is above a threshold (Prob.ind.overexpo.am.overX, default 20%)
Customizable parameters
Probability for credible intervals (default 90%)
Exceedance threshold (default 5%)
Critical percentile (default 95%)
Threshold for the within-worker correlation coefficient (default 0.2)
Coverage of the population for the R ratio (default 80%)
Threshold for the probability of individual overexposure (default 20%)

* *In addition: for any individual worker: all metrics from the SEG analysis*

Finally, while estimating the lognormal probability distribution is at the heart of industrial hygiene data interpretation, the normal distribution is also used in some cases (e.g., while chemical exposures, the focus of this report, most often follow the lognormal model, noise exposure levels expressed in decibels are usually normal). Moreover, both distributions are closely related, since if X follows a normal distribution with GM and GSD, Y=ln(X) follows a normal distribution with mean ln(GM) and standard deviation ln(GSD). For that reason, all Bayesian models were straightforward to adapt to the normal case, and therefore have the option to analyse the data either as lognormally (the default option) or as normally distributed. In the following sections, the main focus is on the lognormal model, but a subsection shortly describes the normal option and its particularities. Results specifically associated with the normal models are presented in Appendix C. Table C1 in Appendix C summarizes the metrics calculated in WebExpo for the normal model.

4.2 Specific objective 2: Creation of a library of computer code

Detailed mathematical presentation of the models and MCMC algorithms set up by the McGill team is available in Appendix B.

4.2.1 Bayesian models created in WebExpo - SEG analysis

The main assumption underpinning this model is that the exposure regimen studied is well represented by a lognormal distribution, and that a representative sample of that distribution has been obtained.

Let X be the random variable representing exposure levels.

Let Y be defined as Y=ln(X). Y therefore corresponds to the log transformed exposure levels.

Since X follows a lognormal distribution, Y follows a normal distribution, that can be expressed as $Y \sim N(\mu,\sigma)$.

The geometric mean of the exposure distribution is defined by $GM=\exp(\mu)$.

The geometric standard deviation is defined by $GSD=\exp(\sigma)$.

μ and σ represent the unknown parameters of the model.

4.2.1.1 Definition of prior distributions

As μ and σ are the parameters of interest in this model, we needed to set up prior distributions for both parameters.

For our primary model [SEG.informedvar, section 3 of Appendix B], we selected a weakly informative prior distribution for μ, in the form of a bounded uniform distribution as described, e.g., in Huynh *et al.* or Banerjee *et al.* (S. Banerjee *et al.*, 2014; Huynh *et al.*, 2016). For σ, we took inspiration from the model described by McNally *et al.* (McNally *et al.*, 2014), in which they used the population of values observed in a dataset presented by Kromhout *et al.* (Kromhout *et al.*, 1993) and Rappaport *et al.* (Rappaport *et al.*, 1993). The authors described variability estimates for close to 200 exposure groups. From Table A1 in Kromhout *et al.* we tabulated 165 values of σ. Graphical assessment suggested a lognormal shape for the distribution of these

values. Fitting the data to a lognormal distribution yields a GM of 0.84 for σ (which corresponds to a GSD of 2.32 for exposure levels), and a GSD of 1.87 (this quantity expresses variability in the sigma values, not in exposure levels). This distribution corresponds to 95% of exposure level GSD values between 1.3 and 17.6. Seventy percent of the distribution is comprised between 1.5 and 4.5. The prior for variability in the [SEG.informedvar] model is therefore expressed as a lognormal distribution for the log-transformed GSD of the exposure distribution. The default values are those above, but can be user-specified.

This choice of prior keeps the level of prior information very low for the geometric mean, but somewhat informative for variability based on historical data. With flexibility in mind, we added two additional choices of priors for this model:

1- [SEG.uninformative, section 2 of Appendix B]: This model has a uniform prior for σ as well as μ, with the ranges selected by the user. By selecting wide ranges, the model becomes uninformative in practice.

2- [SEG.riskband, section 5 of Appendix B]: This model expands the proposition by Hewett *et al.*, where the prior information involves setting upper and lower bounds for σ as well as μ, but also assigning probabilities for the 95th percentile of the exposure management bands defined by AIHA (Jahn *et al.*, 2015). These bands are defined by: <0.01*OEL, [0.01*OEL-0.1*OEL], [0.1*OEL-0.5*OEL], [0.5*OEL-OEL], and ≥ OEL. The last band corresponds to unacceptable exposure. With this prior, the user has to enter a probability for each of these categories, summing to one. Setting all five probabilities to 0.2 corresponds to an uninformative prior (inasmuch as the ranges for μ and σ are reasonably wide), while assigning a high probability to one of the bands will render the prior increasingly informative. Assignment of probabilities can be based on expert judgment, mathematical emission models or other datasets (Arnold, Stenzel, Drolet, & Ramachandran, 2016; Jayjock, Chaisson, Franklin, Arnold, & Price, 2009; P. Logan *et al.*, 2009; P. W. Logan *et al.*, 2011). Aiming at flexibility, we expanded Hewett *et al.*'s proposal to a customizable number of bands and band limits.

In addition, we were interested in allowing users to inform calculations using other relevant data, which would be available in the form of summary parameters (i.e., mean, standard deviation and sample size). Quick *et al.* described such a prior in the *Annals of Occupational Hygiene* after the WebExpo project had started, so incorporating their proposal was not possible without additional resources (Quick *et al.*, 2017). However we proposed a modification of the [SEG.informedvar] model which would create similar results. The mathematics of the proposal is described in Appendix B. In essence the calculations are equivalent to the user submitting to the [SEG.informedvar] model a dataset that would include all data, current observations plus the additional summarised dataset. Users selecting this option have to provide mean (in the same scale as μ), standard deviation (in the same scale as σ), and sample size. The corresponding model is called [SEG.past.data].

4.2.1.2 Censored data analysis

One way to model censored data in Bayesian analysis is to treat them as missing data that are constrained to fall in the censored range of the distribution (Gelman, 2013). Hence, a left-censored data below the level of quantitation (< LOQ) is treated as a missing observation from the part of the distribution that is below the LOQ. At each iteration of the MCMC process, the

missing values are imputed with the corresponding constraint. This constraint influences the posterior distribution of the mean and standard deviation estimated by the model. When the priors have low information, the procedure is close to the frequentist maximum likelihood. We implemented this approach for left-censored, interval-censored, and right-censored data (the last two cases, although not prioritary according to the expert group, are straightforward extensions of the first). In addition, the censoring points can be specific to each observation (i.e., multiple LOQ values are permitted).

4.2.1.3 Measurement error

4.2.1.3.1 Measurement error expressed as a standard deviation

The classical measurement error model for a measured quantity would typically be expressed as the following:

Observed_X=True_X + error

Assuming no bias but only random fluctuations around the true value, the traditional model for the “error” quantity is a normal distribution with mean zero and a fixed, potentially unknown standard error σ_e.This measurement error structure was added as an option in the SEG analysis models.

The Bayesian model would therefore be defined as follows:

$$\textbf{Observed\_X} \sim \textbf{N(True\_X, } \sigma_e\textbf{)} \quad (1)$$

and with True_Y=ln(True_X)

$$\textbf{True\_Y} \sim \textbf{N}(\mu,\sigma) \quad (2)$$

σ_e, is treated as unknown with a bounded uniform prior distribution. If σ_e is assumed known, the user can set the lower and upper bounds as equal.

In practice, the actual mathematical statement of the model in Appendix B is slightly different from above. Hence, as the true values are assumed to follow a lognormal model, they are strictly positive. However, equation 1 may lead to negative values if the standard deviation is large compared to the true value. As a consequence, the normal distribution defining the observed value as a function of the true value is truncated so that only positive observed values can be generated (see section 6.1.1 of Appendix B).

4.2.1.3.2 Measurement error expressed as coefficient of variation

As mentioned previously, it is common in industrial hygiene to express measurement error in terms of CV, i.e., the error is proportional to the exposure level. Typical CVs range from a few percents for an 8 h time weighted average value based on the chemical analysis of an adsorbent tube to ~30% for instantaneous colorimetric detector tubes. A constant CV across a set of measurements implies a different standard deviation for each measurement. Such a model was also created for WebExpo and represents a second measurement error treatment option in WebExpo.

$$\mathbf{Observed\_X \sim N(True\_X,\ CV_e * True\_X)} \quad \mathbf{(3)}$$

with CV_e the coefficient of variation expressing measurement error,

and with True_Y=ln(True_X)

$$\mathbf{True\_Y \sim N(\mu,\sigma)} \quad \mathbf{(4)}$$

In the WebExpo models, as for σ_e, CV_e is treated as unknown with a bounded uniform prior distribution. If CV_e is assumed known without uncertainty, the user can set the lower and upper bounds as equal.

4.2.1.4 Modification of the models for the normal distribution

When the option for the normal distribution is selected, there is no prior log-transformation of the observations. Therefore the parameters μ and σ are directly the mean and standard deviation of the underlying distribution.

Although the normal model accommodates negative values, the WebExpo model for the normal distribution is restricted to positive values. Hence, as the measurement error can be expressed as a CV, negative values of X would imply negative values for the standard deviation of the error term. Hence users interested in fitting data involving negative or near 0 values should transform their data by adding a positive constant prior to analysis.

4.2.1.5 Interpretation of the Bayesian model outputs

As mentioned in section 4.1, the typical output of Bayesian analysis estimated through MCMC is a large sample from the joint posterior distribution of the unknown parameters, from which all inferences are made. Thus in the case of the SEG analysis, the raw output of the algorithms available to users is, e.g., 25 000 μ/σ couples. From these, we applied a number of equations to estimate the metrics described in section 4.1.4.

Figure 3 illustrates the data process flow for the analysis of the lognormal distribution. Inputs include actual observations, occupational exposure limit, parameters specific to the Bayesian model (choice and specification of prior, MCMC parameters, choice and specification of measurement error) as well as parameters used to interpret the samples from the posterior distribution. For the lognormal model, we opted to have the observations divided by the OEL prior to being processed by the Bayesian routines. This standardization ensures that the quantities processed by the Bayesian analyses, whatever the initial unit or state, will be in a range approximately centered on 1. This uniformity allows proposing lower and upper bounds for μ that are adequate for most situations.

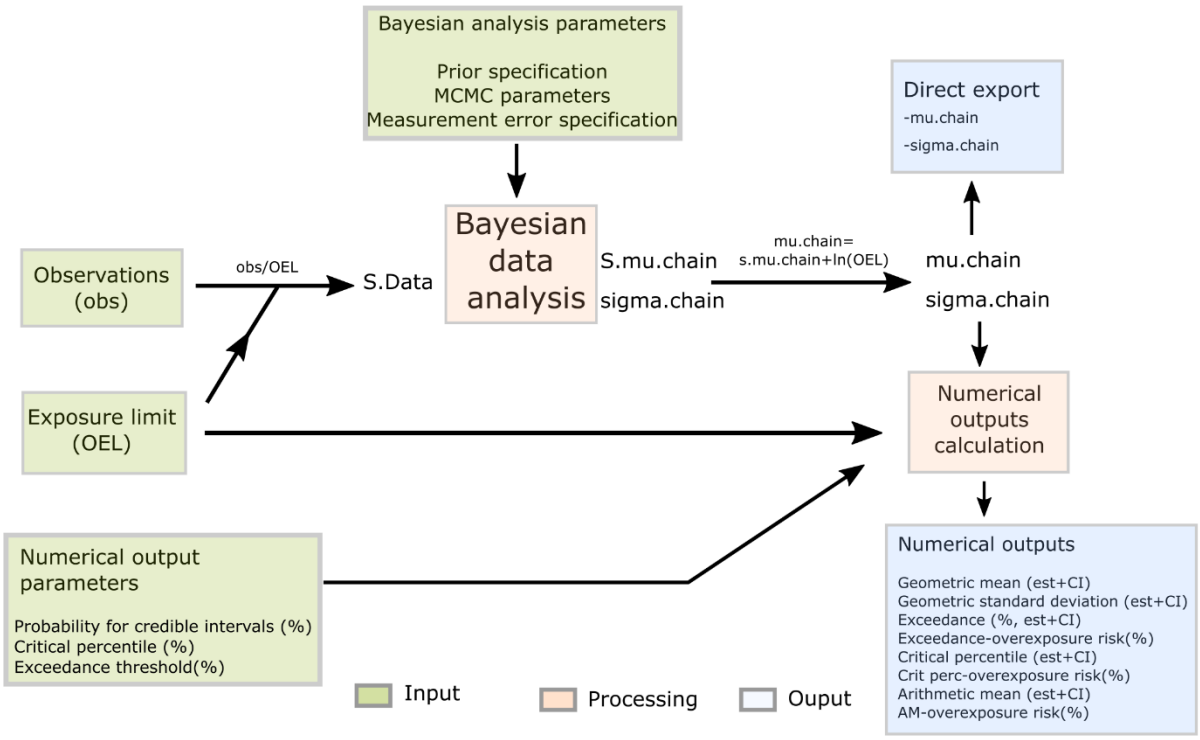


Figure 3. Data processing flow for the SEG analyses – Lognormal distribution.

The output metrics include the geometric mean, the geometric standard deviation, the exceedance fraction of the OEL, any percentile of the distribution (default 95%), as well as the arithmetic mean, obtained from the equations below.

Geometric mean of the exposure distribution:

$$\boldsymbol{GM} = \mathbf{exp}(\boldsymbol{\mu}) \tag{5}$$

Geometric standard deviation of the exposure distribution:

$$\boldsymbol{GSD} = \mathbf{exp}(\boldsymbol{\sigma}) \tag{6}$$

X^{th} percentile of the exposure distribution:

$$\boldsymbol{PX} = \mathbf{exp}\{\boldsymbol{\mu} + \boldsymbol{\Phi}^{-1}(\boldsymbol{X}) * \boldsymbol{\sigma}\} \tag{7}$$

where Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution.

Exceedance fraction of the OEL:

$$\boldsymbol{F(\%) = 100 * \left\{1 - \Phi\left(\frac{\ln(OEL) - \mu}{\sigma}\right)\right\}} \quad \textbf{(8)}$$

where Φ is the cumulative distribution function of the standard normal distribution.

Arithmetic mean of the exposure distribution:

$$\boldsymbol{AM = \exp\{\mu + 0.5 * \sigma^2\}} \quad \textbf{(9)}$$

Uncertainty around the previous metrics is characterised by calculating them for all joint values of μ and σ in the posterior sample. For example, equation 8, applied to the joint posterior sample for μ and σ, will yield 25 000 values of exceedance fraction, which represents our knowledge about this parameter, given the model, the prior, and the observations. These values define the uncertainty surrounding the estimation process. The point estimate for exceedance fraction will be the median of the 25 000 values, and, e.g., their 5th and 95th percentiles will form a 90% equal-tail credible interval. Uncertainty can also be expressed in the form of what we defined as overexposure risk in section 4.1.3: the proportion of the 25 000 posterior values over the threshold of 5% represents the probability that the true exceedance fraction is ≥5%.

4.2.1.6 Examples

Let us illustrate the SEG analysis calculations using a hypothetical dataset coming from a known distribution, with geometric mean True_GM=30, and geometric standard deviation True_GSD=2. The true 95th percentile of this distribution is 84, and its true arithmetic mean is 38. With an arbitrary OEL at 100, the actual exposure regimen would therefore be acceptable according to current consensual overexposure definitions.

We will use a random sample of size nine (recommended in the recent European community guideline) from this true distribution to apply the Bayesian models created for WebExpo. These numbers might represent, for example, nine time-weighted averaged toluene concentrations measured for a SEG.

24.7 / 64.1 / 13.8 / 43.7 / 19.9 / 133 / 32.1 / 15 / 53.7

[sample.1 in Appendix E]

For this example, we will first assume no measurement error and will run the calculations with the [SEG.informedvar] model. The raw output of the Bayesian calculations includes a sample of 25 000 values from the joint posterior distribution of μ and σ, such that μ=ln(True_GM) and σ=ln(True_GSD). The model was run in practice using the R+JAGS algorithms (see 4.3), using the default parameters (see Appendix D).

Figure 4 shows the histograms of the posterior samples of μ and σ. These histrograms reflect the knowledge we gained about μ and sigma given the data, the priors and the model. They represent our estimate of the uncertainty about these parameters. In this example, the most plausible values for μ are probably between 3 and 4 (although as seen on the histogram, more

extreme values are possible). The median of the 25 000 values of the posterior sample for μ in the histogram represents the point estimate for μ: 3.53. Plausible values for sigma would be between 0.5 and 1.5, with a point estimate of 0.78.

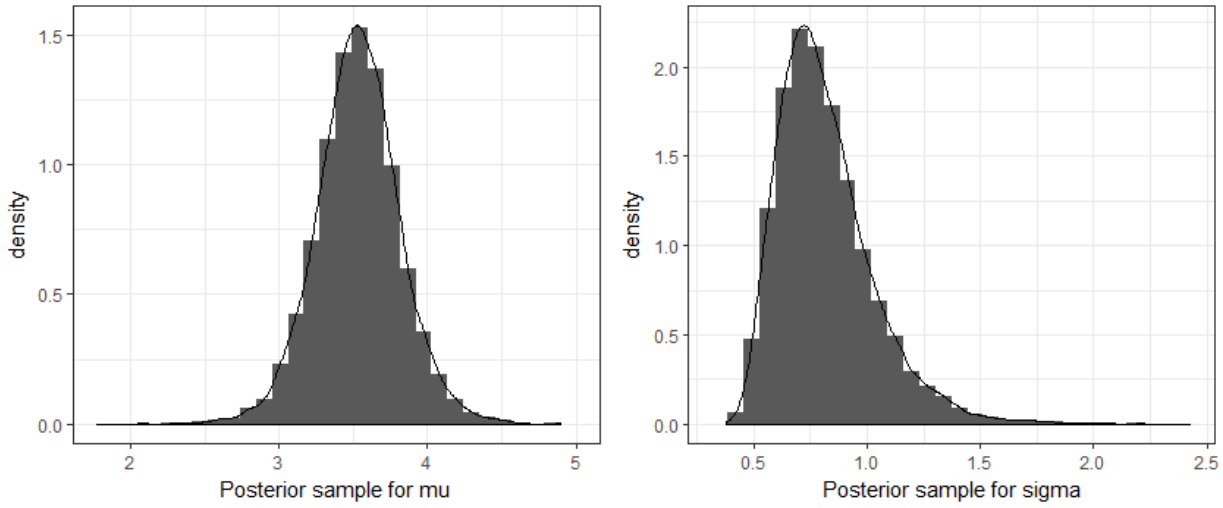


Figure 4. Posterior samples for the log-transformed geometric mean and standard deviation output from the SEG.informedvar model (lognormal model).

From the values illustrated in Figure 4, it is straightforward to apply equations 5-9 to obtain posterior samples for the various metrics of interest. Figure 5 below shows the posterior samples for the 95th percentile of the distribution (eq. 7) and the arithmetic mean of the distribution (eq. 9).

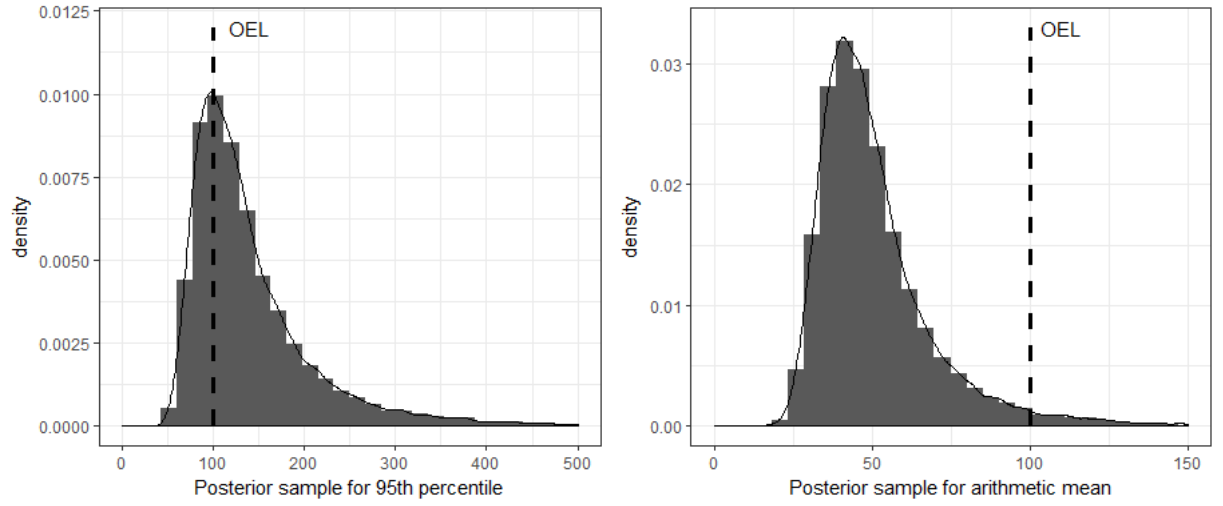


Figure 5. Posterior sample for the 95th percentile and arithmetic mean, calculated from the output from the SEG.informedvar model (lognormal model).

The histograms shown in Figure 5 illustrate the uncertainty surrounding the estimation of these metrics from a sample of size 9, since the values in the histograms cover a wide range. The notion of overexposure risk is well illustrated in Figure 5: using the criterion 95th percentile

≥ OEL (e.g., OEL=100 µg/m^3) for overexposure, the probability (or risk) of overexposure is represented by the proportion of the area in the histogram to the right of the vertical bar showing the OEL, i.e., the proportion of values in the posterior sample of the 95th percentile that are above the OEL.

Table 3 summarizes the results of the interpretation of the posterior samples, including point estimates and 90% credible intervals, as well as overexposure risk (95th percentile and AM), and the AIHA risk band probabilities, i.e., respective probabilities that the true 95th percentile (or AM) is <0.01*OEL, [0.01*OEL-0.1*OEL], [0.1*OEL-0.5*OEL], [0.5*OEL-OEL], and ≥ OEL.

To illustrate the notion of point estimate and credible interval, Table 3 indicates that the most plausible value for GM is 34.2, with a 90% probability that the actual value is between 21.6 and 54.5. Table 3 also indicates that the most plausible value for exceedance fraction is 8.29%, with a 90% probability that the actual value is between 1.46 and 26.5%, and a 71% probability that it is ≥5% (overexposure risk). In terms of the AIHA risk bands for the 95th percentile, while the point estimate is 122 compared to the OEL of 100, Table 3 shows that there is a 71% probability that the actual 95th percentile is >OEL (100 µg/m^3), 29% chance that it is between 0.5*OEL (50 µg/m^3) and OEL (100 µg/m^3), and <1% probability that it is in the other categories. Thus, despite a true 95th percentile below the OEL, inference from the available sample suggests there is a high probability (71%) that the true value (which we know is < OEL) is above the OEL. This illustrates the difficulty to make conclusive inference based on limited sample size when the true situation is acceptable, but marginally so. Table 3 also illustrates the difference between overexposure being defined based on the 95th percentile (overexposure risk 71%) vs. the arithmetic mean (overexposure risk 3.7%).

Table 3. Exposure metrics point estimates and credible intervals for an example of Bayesian calculation for the lognormal model

Parameter	**Point estimates and 90% credible interval**
GM	34.2 [21.6 - 54.5]
GSD	2.18 [1.72 - 3.37]
Exceedance fraction (%)	8.29 [1.46 - 26.5] Overexposure risk: 71%
95th percentile	122 [72.1 - 303] Overexposure risk: 71%
AIHA band probabilities in % (95th percentile)	0 / 0 / 0.048 / 29 / 71
Arithmetic mean	46.7 [30.4 - 91.6] Overexposure risk: 3.7%
AIHA band probabilities in % (AM)	0 / 0 / 59 / 37 / 3.7

In order to illustrate the influence of the choice of prior distribution, we analysed the same sample using the other WebExpo options for the prior information. We included the uninformative model, the riskband model (similar to the proposal by Banerjee *et al.*, and Hewett *et al.*), as well as the past.data model. For the informative riskband model, we defined prior knowledge as a prior assessment from a hypothetical expert judging the situation likely acceptable albeit not by a great margin, therefore choosing the following prior probabilities for the AIHA bands: <0.01*OEL (10%), [0.01*OEL-0.1*OEL] (20%), [0.1*OEL-0.5*OEL] (50%),

[0.5*OEL-OEL] (10%), and ≥ OEL (10%). For the past.data model, we will consider the prior existence of a dataset of five measurements, with geometric mean 5 and geometric standard deviation 2.4, judged relevant for the current analysis. For these analyses, parameters other than those mentioned above were the default parameters described in Appendix D. Table 4 shows the results.

Table 4. Exposure metrics point estimates and credible intervals for 4 choices of prior distribution

Parameter	Informedvar	Uninformative	Past.data	Riskband
GM (90% CrI)	34.2 [21.7 - 54.1]	34.3 [20.9 - 56.8]	17.2 [9.91 - 29.7]	29.8 [19.1 - 46.1]
GSD (90% CrI)	2.18 [1.73 - 3.38]	2.3 [1.75 - 4.15]	3.33 [2.49 - 5.45]	2 [1.66 - 3.19]
Exceedance fraction (%) (90% CrI)	8.30 [1.51 - 26.3]	9.77 [1.76 - 30.3]	7.16 [1.81 - 19.7]	3.71 [0.872 - 21.2]
95^{th} percentile (90% CrI)	122 [72.8 - 302]	134 [74.9 - 418]	124 [64 - 342]	90.8 [65.6 - 247]
Overexposure risk (%, P95)	71%	76%	69%	26%
AM (90% CrI)	46.6 [30.7 - 91.3]	49.1 [31.2 - 118]	35.9 [20.2 - 90.4]	37.6 [27.3 - 76.6]
Overexposure risk (%, AM)	3.7%	7.5%	3.7%	2.4%

Table 4 illustrates the influence of different choices of prior distribution when the sample size is relatively small. While both weakly informative priors (informedvar and uninformative) yield similar (albeit not equal) results, both informative priors have a marked influence. Hence the past.data procedure, involving a dataset with lower levels than those in the sample, decreased the estimate of GM, while increasing the estimate for GSD (probably due to the discrepancy in levels between the two datasets). The combined decrease in GM and increase in GSD resulted overall in relatively little change in the exceedance fraction and 95^{th} percentile, but caused a decrease in the arithmetic mean. The riskband prior, which represented lower exposures, pulled the estimates towards lower values, with a decrease in GM leading to lower 95^{th} percentile, exceedance fraction, arithmetic mean, and associated overexposure risk values.

In order to illustrate the impact of measurement error on the analysis of exposure measurements, we present the analysis of a sample coming from a known distribution. We first generated a sample of size 100 from a lognormal distribution with GM=60 and GSD=1.5 [sample.2 in Appendix E]. Measurement error was added in the form of a random deviation from the true underlying exposure value, for each point, characterized by a coefficient of variation of 30%. We then analysed this sample using three approaches: an analysis assuming no measurement error, an analysis assuming that measurement error is known to be 30%, and an analysis where measurement error is unknown but supposed between 15 and 45%. Table 5 shows the result of this analysis.

Table 5. Exposure metrics point estimates and credible intervals in the presence of measurement error

Parameter	No measurement error (A)	Known CV (30%)	Unknown CV (15-45%)
GM (90% CrI)	56.9 [52.2 - 61.9]	59.8 [54.8 - 65.1]	58.9 [53.7 - 64.3]
GSD (90% CrI)	1.68 [1.59 - 1.79]	1.49 [1.39 - 1.62]	1.55 [1.4 - 1.7]
Exceedance fraction (%) (90% CrI)	13.7 [9.62 - 18.8]	9.8 [5.27 - 15.8]	11.2 [5.82 - 17.1]
95^{th} percentile (90% CrI)	133 [118 - 153]	115 [101 - 135]	121 [103 - 142]
Arithmetic mean (90% CrI)	65 [59.6 - 71.5]	64.8 [59.4 - 70.9]	64.8 [59.3 - 71]

(A): The data actually contain measurement error; the column heading indicates the type of analysis that was applied to these data.

Table 5 shows that not taking the measurement error into account caused little effect on the GM, while it caused an overestimation of the GSD. Overestimating GSD impacted the estimation of the upper tail of the distribution with an associated overestimation of the exceedance fraction and of the 95^{th} percentile. The impact was lower for the arithmetic mean. It is noteworthy that the credible interval for GSD for the naïve analysis didn't include the true value, as opposed to the two analyses including measurement error. Compared to the analysis assuming a known CV, assuming a CV known only within a range caused little effect for that example.

4.2.2 Bayesian models created in WebExpo - Between-worker difference analysis

The main assumption underpinning this model is that the exposure regimen within an exposure group is adequately represented by the following hierarchical structure: workers within their group have their personal exposure distribution adequately represented by a lognormal distribution. The workers' distributions differ in their location (GM), but not in their variability. The collection of worker-specific GMs themselves follows a lognormal distribution.

Let X be a random variable representing exposure levels.

Let Y=ln(X) where Y then corresponds to the log -transformed exposure levels.

Let y_{ij} be the value corresponding to the measurement taken on the j^{th} day for the i^{th} person.

The one level hierarchical random effects model is written as:

$$y_{ij}=\mu_y+b_i+e_{ij}$$

for i=1,2,…,k workers on j=1,2,…,n_i days

μ_y is the group mean, b_i is the random effect for worker i, and e_{ij} is the random deviation on the j^{th} day from the i^{th} worker's mean μ_y+b_i.

Under this random effect model, b_i and e_{ij} are mutually independent and normally distributed with means of zero. The between-worker standard deviation (of b_i) is σ_b, and the within-worker standard deviation (of e_{ij}) is σ_w.

The group geometric mean is defined by $GM=\exp(\mu_y)$.

The between-worker geometric standard deviation is defined by $GSD_B=\exp(\sigma_b)$.

The within-worker geometric standard deviation is defined by $GSD_W=\exp(\sigma_w)$.

Any individual worker's exposure distribution is defined by:

$GM_i=\exp(\mu_y+b_i)$ and

$GSD_i=\exp(\sigma_w)$.

4.2.2.1 Definition of prior distributions

For this model (described in section 4 of Appendix B), we needed to set up prior distribution for the three parameters μ_y, σ_b, and σ_w.

For our primary analysis, we set up priors similar to the [SEG.informedvar] model described above. The prior information for μ_y is the same as for the SEG model, a uniform distribution bounded by -20 and 20.

For the variability parameters, we used the same published data as for the [SEG.informedvar] model, but used the between- and within-worker components of variance presented in the Kromhout *et al.*'s paper, as opposed to total variability. Graphical assessment also suggested a lognormal shape for the distribution of between-worker (σ_b) and within-worker (σ_w) standard deviations. Fitting the data to lognormal distributions yielded the following parameters:

For between-worker variability:

GM of the lognormal distribution for σ_b : 0.415
GSD of the lognormal distribution for σ_b : 2.18
This distribution corresponds to 95% GSD_B values between 1.1 and 6.7. Seventy percent of the distribution is comprised between 1.2 and 2.5.

For within-worker variability:

GM of the lognormal distribution for σ_w : 0.844
GSD of the lognormal distribution for σ_w : 1.88
This distribution corresponds to 95% GSD_W values between 1.3 and 18.3. Seventy percent of the distribution is comprised between 1.6 and 5.1.

This choice of prior keeps the level of prior information very low for the geometric mean, but somewhat informative for variability based on historical data. It is very similar to those described by McNally *et al.* (McNally *et al.*, 2014). As in the case of the SEG analysis, the above distributional parameters for the priors are proposed default values, but can be user-selected.

We added one additional choice of prior for this model, where the prior distributions for σ_b and σ_w are uniform, with user-selected bounds. Large bounds would correspond to an uninformative Bayesian prior.

4.2.2.2 Censored data analysis

Censored data in the between-worker differences model is processed using the same approach described in 4.2.1.2. Multiple censoring points are permitted, and data can be left, right, or interval censored. The mathematics are detailed in Appendix B.

4.2.2.3 Measurement error

Measurement error for this model is treated the same was as described in 4.2.1.3 and detailed in Appendix B.

4.2.2.4 Modification of the models for the normal distribution

The changes to the Bayesian models for the normal instead of lognormal model for between-worker differences analysis are the same as described in 4.2.1.4.

4.2.2.5 Interpretation of the Bayesian model outputs

In the case of the between-worker difference analysis, the raw output of the algorithms available to users is, e.g., 25 000 joint μ_y / σ_b / σ_w / b_i (i=1 to k workers) values. From these, we applied several equations to estimate the metrics described in section 4.1.4 (see below).

Figure 6 below illustrates the data process flow in the analysis of the lognormal distribution. Inputs include the actual observations with a worker identifier, the occupational exposure limit, parameters specific to the Bayesian model (choice and specification of prior, MCMC parameters, choice and specification of measurement error) as well as parameters used to interpret the samples from the posterior distribution. As for the SEG analysis of the lognormal model, we opted to have the observations divided by the OEL prior to the Bayesian routines.

The following equations describe how the various metrics presented in 4.1.4 are calculated from the MCMC output.

Group geometric mean:

$$\boldsymbol{GM_{group} = \exp(\mu_Y)} \quad \textbf{(10)}$$

Between-worker geometric standard deviation:

$$\boldsymbol{GSD_b = \exp(\sigma_b)} \quad \textbf{(11)}$$

Within-worker geometric standard deviation:

$$\boldsymbol{GSD_w = \exp(\sigma_w)} \quad \textbf{(12)}$$

Within-worker correlation coefficient:

$$\boldsymbol{\rho = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2}} \quad \textbf{(13)}$$

$R_{X\%}$ ratio : Fold range containing the middle X% of the distribution of either worker specific geometric means, arithmetic mean or any percentile. Initially X was set at 95% by the first proponents of its use (Kromhout et al., 1993; Rappaport et al., 1993). Our proposed default value is 80%.

$$\boldsymbol{R_{X\%} = \exp(2 * \Phi^{-1}\left(\frac{1+X}{2}\right) * \sigma_b)} \quad \textbf{(14)}$$

Probability that a single random worker would have his own arithmetic mean above the OEL

$$\boldsymbol{P_{ind}^{AM}(\%) = 100 * \left\{1 - \Phi\left(\frac{\ln(OEL) - \left(\mu_Y + 0.5 * \sigma_w^2\right)}{\sigma_b}\right)\right\}} \quad \textbf{(15)}$$

Probability that a single random worker would have his own Xth percentile above the OEL (this is equivalent to the probability that a single random worker would have his own exceedance of the OEL above (100-X)%:

$$\boldsymbol{P_{ind}^{PX}(\%) = 100 * \left\{1 - \Phi\left(\frac{\ln(OEL) - \left(\mu_Y + \Phi^{-1}(X) * \sigma_w\right)}{\sigma_b}\right)\right\}} \quad \textbf{(16)}$$

In addition to the above, it is also possible to obtain metrics specific to any individual exposure distribution. Hence by definition the exposure distribution for worker *i* is defined by:

Geometric mean of the exposure distribution:

$$\boldsymbol{GM = \exp(\mu_Y + b_i)} \quad \textbf{(17)}$$

Geometric standard deviation of the exposure distribution:

$$\boldsymbol{GSD = \exp(\sigma_w)} \quad \textbf{(18)}$$

Xth percentile of the exposure distribution:

$$\boldsymbol{PX = \exp\{\mu_Y + b_i + \Phi^{-1}(X) * \sigma_w\}} \quad \textbf{(19)}$$

where Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution.

Exceedance fraction of the OEL:

$$\boldsymbol{F(\%) = 100 * \left\{1 - \Phi\left(\frac{\ln(OEL) - \mu_Y - b_i}{\sigma_w}\right)\right\}} \quad \textbf{(20)}$$

where Φ is the cumulative distribution function of the standard normal distribution.

Arithmetic mean of the exposure distribution:

$$\boldsymbol{AM = \exp\{\mu_Y + b_i + 0.5 * \sigma_w^2\}} \quad \textbf{(21)}$$

It should be noted that the above worker-specific metrics, despite being applicable to a single worker, are estimated through fitting the Bayesian model to the entire dataset, not just data from worker *i*.

As for the SEG analysis, uncertainty around the previous metrics is characterised by calculating them for all joint values of μ, σ_B, σ_w, and the b_i values in the posterior sample.

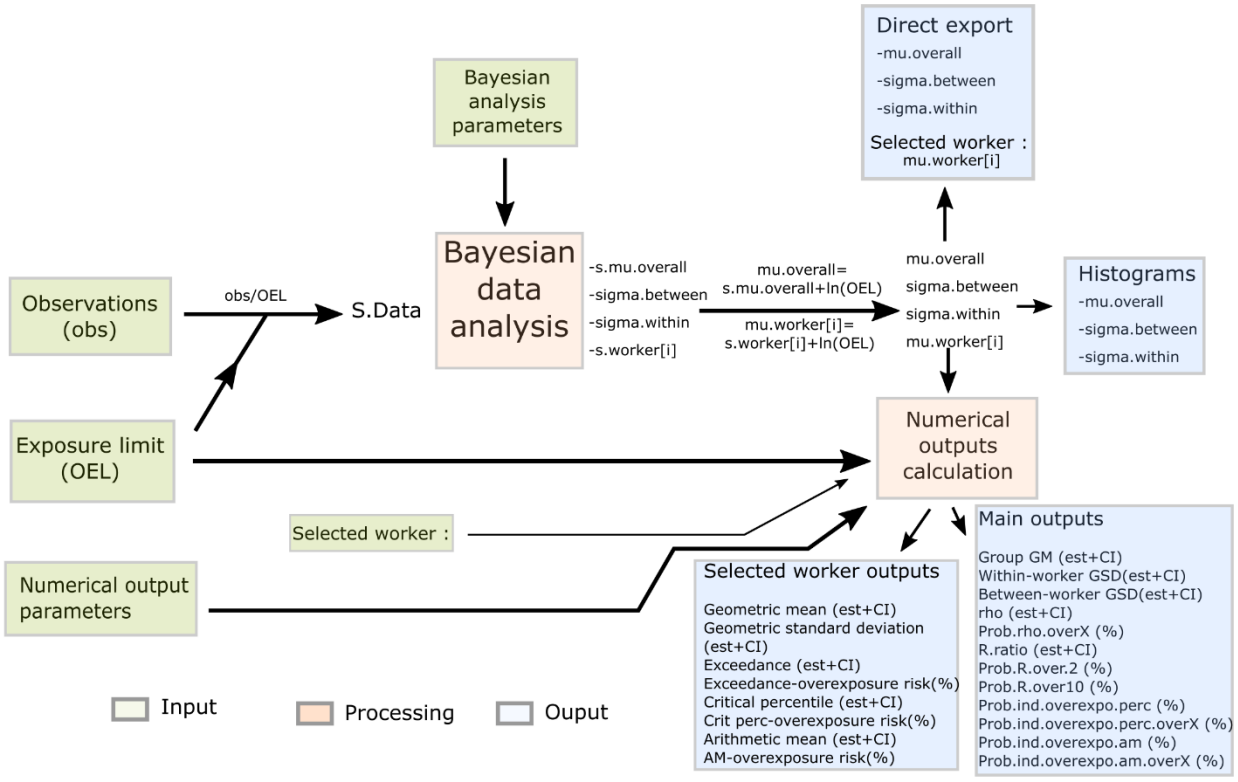


Figure 6. Data processing flow for the between-worker difference analyses – Lognormal distribution.

Examples

We shall illustrate the between-worker analysis with the analysis of two fictitious samples drawn from known distributions. The first distribution has a group GM of 30, with an overall GSD of 2.5, and has low within-worker correlation: rho=0.06 (this value is the 25th percentile of the distribution of rho values in the Kromhout *et al.* dataset described in 4.2.2.1 and 4.2.1.1). The second distribution has the same group GM and GSD but has high within-worker correlation: rho=0.66 (this value is the 75th percentile of the distribution of rho values in the Kromhout *et al.* dataset).

From each distribution we drew a sample of 100 observations: ten observations from ten different workers. We shall perform the analysis considering an OEL of 150, which is slightly higher than the group theoretical 95th percentile (true group P95=135). The two samples are presented in Appendix E [sample.3 and sample.4, respectively].

For this example, we assumed no measurement error and ran the calculations with the [between-worker differences.informedvar] model implemented in R + RJAGS (see 4.3).

The raw output of the Bayesian calculations includes a sample of 50 000 values from the joint posterior distribution of μ_y / σ_b / σ_w / b_i (i=1 to k workers). The interpretation of these values is similar to the one described in 4.2.1.6. Table 6 summarizes the results of the interpretation of the posterior samples, including point estimates and 90% credible intervals, as well as overexposure risk metrics.

Table 6. Exposure metrics point estimates and credible intervals for an example of Bayesian calculation for the lognormal model (between-worker difference analyses)

Parameter	Low within-worker correlation (rho=0.06)	High within-worker correlation (rho=0.66)
Group GM (90% CrI)	28.6 [23.6 - 34.7]	28.3 [18.7 - 43.4]
Between-worker GSD (90% CrI)	1.24 [1.09 - 1.54]	2.15 [1.72 - 3.18]
Within-worker GSD (90% CrI)	2.34 [2.14 - 2.62]	1.74 [1.64 - 1.88]
Within-worker correlation (rho) (90% CrI)	0.06 [0.00908 - 0.206]	0.654 [0.47 - 0.818]
Probability that rho>0.2	5.5%	100%
R.ratio (90% CrI)	1.74 [1.24 - 3.01]	7.09 [3.99 - 19.4]
Probability that R>2	30%	100%
Probability that R>10	0%	25%
Probability of individual overexposure (95th percentile) in % (90% CrI)	12.1 [0.0248 - 52.4]	16.3 [4.94 - 36.5]
Chances that the above probability is >20%	34%	36%
Probability of individual overexposure (arithmetic mean) in % (90% CrI)	9.5e-08 [0 - 0.177]	2.38 [0.164 - 13.1]
Chances that the above probability is >20%	0%	1.4%

Table 6 shows that for both samples, the analysis, based on a relatively large sample size, yields results very close to the theoretical distribution in terms of group GM and of the values of rho for the low and high variability samples. The group GSD estimates (not provided in Table 6) are also close to the theoretical value of 2.5: respective point estimates of 2.4 and 2.6 for the low and high within-worker correlation samples.

For the low within-worker correlation sample, the low correlation with a group GSD at 2.5 yields a low between-worker GSD (1.24), illustrated in the low R ratio (1.7), with 70% chances that the true value would be below 2, a threshold initially proposed by Kromhout *et al.* (1993) to define "homogeneous". With low between-worker variability, most of the variability therefore occurs within workers, with a within-worker GSD of 2.3. The opposite phenomenon is observed for the second sample, with a between-worker GSD of 2.15, corresponding to a R ratio of 7, with 100% chances that the true value is above the threshold of 2, and 25% chance that it is above 10. In this case, most of the total variability occurs between workers, with a corresponding low within-worker GSD (1.7).

To illustrate the notion of individual overexposure, for the low within-worker correlation case, Table 6 indicates that the probability that a random worker would have his own 95th percentile above the OEL is estimated to be 12.1% (90% CrI 0.02% - 52.4%). The chances that the true value for this probability is >20%, the criteria used by NvVA and BOHS, are 34%. Similarly, the probability that a random worker would have his own arithmetic mean above the OEL is estimated to be ~0% (90% CrI 0.0% - 0.2%). The chances that the true value for this probability is >20%, the criteria used by NvVA and BOHS, are 0%.

The relative similarity of probability of individual overexposure (both for the 95th percentile, close to 15% or the arithmetic mean, close to 0%) in both samples despite important differences in between-worker variability is noteworthy. This is a consequence of a shared global variability in both groups: for sample one, despite little differences between workers in terms of GM (as measured by the R ratio), the high day-to-day (within-worker) variability implies a potential for relatively elevated values of the worker specific 95th percentile or AM (which both depend on σ_w), similar across workers. For sample two, despite important differences between workers in terms of GM (as measured by the R ratio), the low day-to-day variability implies relatively low values of the worker specific 95th percentile or AM (which both depend on σ_w), but the large differences across workers might cause elevated values for some individuals. To further illustrate this point we present in Table 7 the worker-specific exposure metrics for the lowest and highest exposed (in terms of GM) workers in both samples.

Table 7 clearly shows the important differences between the two samples in terms of the contrast in the GMs for the low and high exposed workers: GM_{least}=26 and GM_{most}=33 for the sample with low within-worker correlation, and GM_{least}=7 and GM_{most}=130 for the sample with high within-worker correlation. Looking at the 95th percentile point estimates shows that workers in the first sample all have similar 95th percentiles, all with (again, not taking uncertainty into account) an acceptable exposure distribution, albeit somewhat marginally (95th percentile around 100-140 for an OEL of 150). For the other sample, the worker-specific 95th percentile estimates vary from 18 (very low compared to the OEL, a clearly acceptable situation) to 325 (more than twice the OEL, a clearly unacceptable situation). These contrasts illustrate the proposition from Kromhout *et al.* and Rappaport *et al.* that using this type of model can be useful

to direct prevention measures towards collective *vs*. focused individual measures (Kromhout *et al.*, 1993; Pesch *et al.*, 2015).

Table 7. Worker specific exposure metrics point estimates and credible intervals for the least and most exposed workers in two samples with low and high within-worker correlation

Parameter	Low within-worker correlation (rho=0.06)		High within-worker correlation (rho=0.66)	
	Least exposed worker (GM)	Most exposed worker (GM)	Least exposed worker (GM)	Most exposed worker (GM)
GM	26.2 [18.7 - 34.7]	33.4 [25.1 - 48.4]	7.13 [5.32 - 9.53]	130 [97.3 - 174]
GSD	2.34 [2.14 - 2.62]	2.34 [2.14 - 2.62]	1.74 [1.64 - 1.88]	1.74 [1.64 - 1.88]
Exceedance fraction (%)	2.02 [0.57 - 5.06]	3.97 [1.57 - 9.53]	0 [0 - 0]	40.1 [22.1 - 60.6]
95^{th} percentile	106 [73.6 - 151]	137 [99.5 - 202]	17.8 [13.1 - 24.7]	325 [243 - 445]
Arithmetic mean	37.8 [26.7 - 51.2]	48.3 [36.1 - 70.1]	8.32 [6.21 - 11.2]	152 [114 - 204]

Table 8 illustrates the analysis of a sample with the same group GM and GSD as above, but with average within-worker correlation (rho=0.22, the median value in the Kromhout *et al.* database), and estimated from a realistic sample size corresponding to the BOHS-NVvA guidelines (n=12, with four repeated measurements on three workers) [sample.5 in Appendix E].

Table 8. Exposure metrics point estimates and credible intervals for an example of Bayesian calculation for the lognormal model (between-worker difference analyses) with realistic sample size

Parameter	Point estimate and 90% credible interval
Group GM (90% CrI)	30.9 [16.9 - 56.4]
Between-worker GSD (90% CrI)	1.4 [1.11 - 2.56]
Within-worker GSD (90% CrI)	2.31 [1.84 - 3.42]
Within-worker correlation (rho) (90% CrI)	0.135 [0.0121 - 0.577]
Probability that rho>0.2	37%
R.ratio (90% CrI)	2.35 [1.29 - 11.1]
Probability that R>2	62%
Probability that R>10	5.8%
Probability of individual overexposure (95^{th} percentile) in % (90% CrI)	29.7 [0.0114 - 99.2]
Chances that the above probability is >20%	59%
Probability of individual overexposure (arithmetic mean) in % (90% CrI)	0.0204 [0 - 21.3]
Chances that the above probability is >20%	5.4%

The results in Table 8 show the important uncertainty in the parameters estimated from the between-worker differences model at current proposed sample sizes. Hence the 90% credible interval for R includes both values considered as showing homogeneous exposure (<2) and very heterogeneous exposure (>10). This is also reflected in the uncertainty around the probability of individual exposure (for the criteria 95th percentile>OEL), estimated for this sample at 29.7%, with a 90% credible interval ranging from 0.01 to 99.2%.

4.3 WebExpo algorithms

The main deliverable of the WebExpo project is the public availability of the algorithms that implement the Bayesian and numerical analyses described in this report. The algorithms are all available from https://github.com/webexpo/ and are licensed under the Appache 2.0 open source licence[15]. This report constitutes the accompanying documentation.

As described in section 3.3.2, the models and numerical interpretation of the posterior samples were first written in the R language. The models themselves were coded in two ways. First, they were written in pure R code, with no call to an external package, to permit translation into C# and JavaScript. As mentioned in 3.2.2.1, they were also coded using the JAGS application to allow R users to perform the calculations with optimal efficiency.

In a second step, the pure R code was translated into C# and JavaScript. During the initial translating effort, it quickly became apparent that the resources required to translate the measurement error component of the pure R code would be very costly, and that their full implementation would be done at the cost of other components of the project (i.e., less prior distribution options). We elected to prioritize translation with the following objectives:

1. To ensure that core SEG and between-worker difference analyses are implemented in all languages for the lognormal and normal models;

2. To ensure that the measurement error is implemented in one language outside R.

Therefore, measurement error was implemented only in C# language, and only with error expressed as a CV (the type thought to be the most useful). The treatment of right/left and interval censored data was included in all languages. Table 9 provides details about the various components implemented in each of the four language settings.

[15] https://www.apache.org/licenses/LICENSE-2.0

Table 9. Various components implemented in each of the four language settings

	No measurement error				Measurement error as CV				Measurement error as SD			
	R	R+JAGS	C#	Java-Script	R	R+JAGS	C#	Java-Script	R	R+JAGS	C#	Java-Script
SEG analysis												
Informedvar	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	-	-
Uninformative	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	-	-
Past.data	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	-	-
Riskband	✓	✓	-	-	✓	✓	-	-	✓	✓	-	-
Between-worker differences												
Informedvar	✓	✓	✓	✓	✓	✓	-	-	✓	✓	-	-
Uninformative	✓	✓	✓	✓	✓	✓	-	-	✓	✓	-	-

4.3.1 Organisation of the scripts

The WebExpo R (pure R and R+JAGS) scripts are organised in the following subsections:

Data formatting scripts: These scripts are used to format the observations prior to submission to the Bayesian calculation functions.

Bayesian model scripts: This represents the major part of the created library. The corresponding algorithms perform the MCMC sampling and their output includes the samples from the posterior distributions.

Bayesian output interpretation scripts: The scripts apply equations 5 to 21 to the posterior samples to obtain the exposure metrics described in the corresponding sections.

In addition, the R scripts include random data generation functions, which generate custom random samples corresponding to the various models implemented in WebExpo. Finally, the R library also includes scripts to replicate all numerical results presented in the present report.

While using the same overall architecture, the C# and JavaScript libraries are organized somewhat differently due to the nature of the programming languages. Like the examples in the R library, the C# and JavaScript libraries also include illustrative algorithms guiding users across the various Bayesian models. Finally, the C# and JavaScript libraries include the code corresponding to the prototypes, comprising a data entry interface and a numerical result output interface.

4.3.2 Calculation parameters

The WebExpo algorithms require an ensemble of inputs in order to perform the MCMC calculations and to provide interpreted exposure metrics as well as the associated uncertainty. They can be separated into three main categories: observations, Bayesian parameters and numerical interpretation parameters.

4.3.2.1 Observations

This category includes the actual values entered by the user. In theory, there is no lower limit to the number of observations supplied by the users, as performing Bayesian calculation without any data will just cause the posterior samples to replicate the prior distributions. We, however, recommend a pragmatic threshold of a minimum of three uncensored observations (e.g., six measurement concentrations including three nondetects). Similarly, there is not a theoretical upper limit to the number of observations submitted for analysis, although at some point memory usage might exceed the computer capacity. As an example, we analysed a 100 000-observation dataset using the R+RJAGS script for the SEG lognormal model on a regular desktop computer in a little over 120 min. In terms of the sample values themselves, the only restriction is that they should be strictly positive, and within the bounds defined for the prior distributions of the Bayesian models. For the lognormal model, we performed a division by the OEL prior to Bayesian calculations, which permitted to limit the range of values actually analysed, and allowed us to propose default “universal” values for the lognormal priors (e.g., bounds for the mean). We also put in this first data input category the occupational exposure limit, which should be expressed in the same scale as the observations.

4.3.2.2 Bayesian parameters

This series of parameters can be further separated into three parts. First the user must select the model for his analysis. This includes the choice between the normal and lognormal distributions, between the SEG and the between-worker difference analyses, and whether measurement error should be considered.

The second set of Bayesian inputs includes the parameters of the selected prior distributions. Table 10 below presents a summary list of these parameters, while the fully detailed list, accompanied by default values and recommended ranges, can be found in Appendix D.

The third set of parameters includes inputs for the MCM sampling procedures. They comprise the number of burn-in iterations (iterations discarded at the end of the sampling, used to let the MCMC procedure attain a stationary distribution), the number of iterations (i.e., size of the posterior sample), as well as initial values for the parameters to be estimated. We recommend using 25 000 iterations with 2 500 burn-in iterations, for the SEG models, and 50 000/5 000 for the between-worker models. The initial values should be set to plausible values for the parameters of interest, within the bounds set by the prior distributions.

Table 10. Parameters defining prior distributions in the WebExpo models

Model	Prior parameters
SEG.informedvar	Bounds for the uniform distribution for μ
	Distributional parameters for the lognormal distribution for σ
SEG.past.data	Bounds for the uniform distribution for μ
	Distributional parameters for the lognormal distribution for σ
	Mean, standard deviation, and sample size of the external dataset
SEG.uninformative	Bounds for the uniform distribution for μ
	Bounds for the uniform distribution for σ
SEG.riskband	Bounds for the uniform distribution for μ
	Bounds for the uniform distribution for σ
	Number of risk bands and associated limits
	Prior probability associated with each band
Between worker.informed var	Bounds for the uniform distribution for μ
	Distributional parameters for the lognormal distributions for σ_w and σ_b
Between worker.uninformative	Bounds for the uniform distribution for μ
	Bounds for the uniform distributions for σ_w and σ_b

Note : for most parameters, the scale depends on the choice of distribution (normal/lognormal),e.g., for SEG.pastdata, the supplied mean is the arithmetic mean of the observations for the normal model, but the log-transformed OEL-standardized geometric mean for the lognormal model.

4.3.2.3 Numerical interpretation parameters

The numerical interpretation parameters have no influence on the Bayesian calculations *per se*; they are used to transform information in the posterior samples into metrics relevant to risk assessment and express uncertainty around their estimation. They include the following:

- Probability for credible intervals (default 90%);
- Exceedance threshold (default 5%): threshold defining an acceptable proportion of exposure levels above the OEL;
- Critical percentile (default 95%): percentile of interest in the exposure distribution;
- Specific to the between-worker difference analyses:
    - Threshold for the within-worker correlation coefficient (default 0.2): the BOHS guidelines recommend performing a detailed assessment of between-worker differences when the point estimate of this coefficient is above 0.2;
    - Coverage of the population for the R ratio (default 80%): the initial proposition from Kromhout *et al.* used 95% (Kromhout *et al.*, 1993). This would correspond approximately to comparing the most and least exposed workers in a population of 100. Our less stringent proposal would correspond to comparing the least and most exposed workers in a population of ten, a figure closer to practical exposure groups in day-to-day IH interventions;

- Threshold for the probability of individual overexposure (default 20%): the BOHS-NVvA guidelines recommend considering an exposure situation as non-compliant when the probability of individual overexposure (calculated considering the 95th percentile) is above 20%. Rappaport *et al.* proposed using a threshold of 10% and considering probability of individual overexposure defined by the arithmetic mean (Rappaport *et al*., 1995*)*.

4.3.3 Performance

4.3.3.1 Numerical accuracy

As mentioned in section 3.2.2.2, we did not perform simulations to evaluate the estimation accuracy of our models and estimation procedures, as they rely on well established literature. However, we concentrated on verifying the reproducibility of the results across the various platforms (R, R+RJAGS, JavaScript, C#).

C#

For testing the SEG analysis models in C#, we used 1 standard sample for each of 24 scenarios defined by combinations of the following characteristics:

- sample sizes 10 or 100
- low variability (GSD=1.5 for lognormal, SD=2 for normal distribution) or high variability (GSD=3.5 for lognormal, SD=15 for normal distribution)
- no censoring, low censoring (15% for n=100, 30% for n=10) or high censoring (50%)
- normal or lognormal

The samples were fit using the C# and R algorithms for each of the following routines: SEG.informedvar, SEG.past.data informed prior and SEG.uninformative.

For the between-worker analysis, a similar approach was used where C# and R were compared for 24 scenarios:

- 5 workers with 3 measurements per worker, 10 workers with 5 measurements per worker and 20 workers with 20 measurements per worker;
- low within-worker correlation (rho=0.2) or high within-worker correlation (rho=0.8);
- no censoring or 50% censoring;
- lognormal or normal.

Finally, for the measurement error module of the SEG.informedvar and SEG.uninformative functions, 12 scenarios were tested:

- sample sizes of 5, 10 or 100;
- no censoring or 60% censoring;
- lognormal or normal.

For the lognormal distribution, measurement error was defined as unknown between 20 and 30%. For the normal distribution, it was defined as unknown between 0.1 and 1.2%.

Across all these experiments, the C# and R differences, calculated at each of the 25 000 iterations for the unkown parameters μ and σ, were on average around 10^{-15}, with maximum values around 10^{-10}. The differences between C# and R were always several orders of magnitude lower than differences observed within the platform (e.g., obtained by repeating an analysis with R or C# with a different random seed).

JavaScript

During the translation from pure R to JavaScript, we compared quantiles of the posterior samples for the unknown parameters. For the SEG.informedvar, SEG.past.data, and SEG.uninformative functions, the following 4 scenarios were tested for a standard sample of size 100: no censoring or 60% censoring; lognormal or normal distribution.

For the between worker informedvar and between worker uninformative functions, the 8 tested samples included: 10 workers with 5 measurements per worker, 20 workers with 20 measurements per worker; no censoring or 60% censoring; lognormal or normal distribution.

Across these experiments, the JavaScript and R differences, calculated for all unknown parameters at each of the 9 percentiles of the MCMC chains, were on average around 10^{-15}, with maximum values around 10^{-10}. Again the differences between JavaScript and R were always several orders of magnitude lower than differences observed within the platform (e.g., obtained by repeating an analysis with R or JavaScript).

R+RJAGS

For differences between R and RJAGS, the following functions were tested using the protocol described in section 3.2.2, for both the lognormal and normal cases: SEG.informedvar with measurement error (specified as a known CV value), SEG.past.data, SEG.uninformative, SEG.riskband, between worker.informedvar and between worker.uninforma tive (with and without measurement error). The results showed that R and R+RJAGS yielded satisfactorily comparable results. As an illustration, Table 11 presents the results of the analysis of a lognormal sample of size 9 (<25.7 / 17.1 / 168 / 85.3 / 66.4 / <49.8 / 33.2 / <24.4 / 38.3 [sample.6 in Appendix E]), with 30% censored data and high variability (true GSD=2.5). The analysis was run 50 times with R and R+RJAGS, and we show the minimum and maximum values observed for ten parameters. We also show the result of one run using C# and one run using JavaScript.

Table 11 provides insight into the variability expected in such analyses and confirms that the variations are of little significance compared with the uncertainty surrounding the point estimates. These variations are expected to be the most important for credible limits which are tails of the posterior samples, as well as for quantities associated with much uncertainty (such as exceedance and the 95th percentile, which involves the tail of the exposure distribution).

Table 11. Comparability of results across platforms

Parameter	R+JAGS (min) (a)	R+JAGS (max) (b)	R (min) (a)	R (max) (a)	C#	JavaScript
GM point estimate	34.1	34.7	34.0	34.6	34.3	34.4
GM 95% LCL	16.3	17.2	16.5	17.0	16.8	16.6
GM 95% UCL	60.1	61.9	60.2	61.6	60.8	61.0
GSD point estimate	2.63	2.69	2.63	2.68	2.66	2.66
GSD 95% LCL	1.88	1.91	1.89	1.90	1.90	1.90
GSD 95% UCL	5.21	5.56	5.21	5.47	5.32	5.34
Exceedance point estimate (%)	13.2	13.5	13.2	13.5	13.3	13.5
Exceedance 95% LCL (%)	3.33	3.53	3.34	3.51	3.41	3.37
Exceedance 95% UCL (%)	32.6	33.9	33.0	34.0	33.5	33.5
Exceedance overexposure risk (%)	88.8	89.9	88.9	89.6	89.1	89.1

(a), minimum values across 50 analyses; (b), maximum value across 50 analyses; LCL, lower confidence limit; UCL, upper confidence limit.

4.3.3.2 Calculation speed

The speed of calculation for the algorithms in this project depends on multiple factors. In Bayesian MCMC analysis, larger sample size as well as the number of iterations will increase calculation time. In addition, measurement error as well as censorship will also increase calculation time, since additional random values must be generated. Outside of these considerations, different platforms might be better suited/optimized for the algorithms, and computer performance is an important determinant. We didn't conduct an extensive survey to evaluate computing time in a vast array of situations. However, for what we believe would cover most situations (n<50 and no measurement error), all platforms used in WebExpo should perform the calculations (involving 25 000 or 50 000 iterations depending on the model) either instantaneously or within a few seconds. Introducing measurement error is the major determinant of calculation speed in our algorithms: the fastest platform is R+RJAGS, with calculation time within a minute, followed by C# (up to 10 minutes), and lastly R, from 30 minutes for small samples to several hours for the more complex between-worker differences model and higher sample sizes.

4.4 Specific objective 3: WebExpo prototypes

The C# and JavaScript prototypes are available from https://github.com/webexpo/. Both prototypes include the WebExpo models as described in Table 9. Hence, compared to the R and R+JAGS, which include all possibilities described in this report, the C# prototype only allows one type of measurement error (expressed as a CV) for the SEG models, and does not include the risk-band model. The JavaScript prototype does not include treatment of measurement error or the risk-band model. Both prototypes provide the same numerical output, including all metrics listed in Table 2, as well as the posterior samples for users wishing to perform different post-MCMC calculations. In the case of the between-worker analysis, both

prototypes also allow selecting a worker and obtaining worker-specific exposure metrics as well as the posterior sample of the worker specific mean. Finally, both prototypes are available in French and English, and include a multilingual infrastructure amenable to translation into other languages. Figures 7 and 8 respectively show the C# and JavaScript user interfaces.

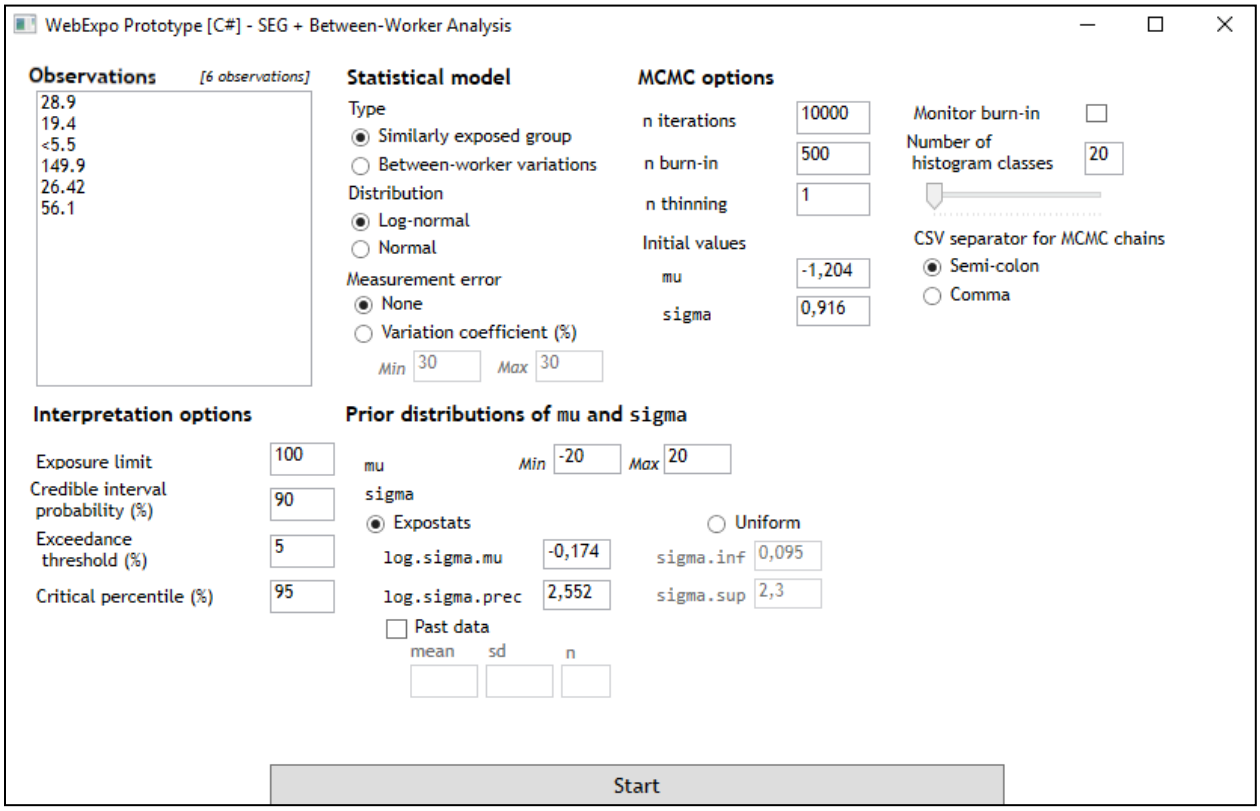


Figure 7. User interface of the C# WebExpo prototype.

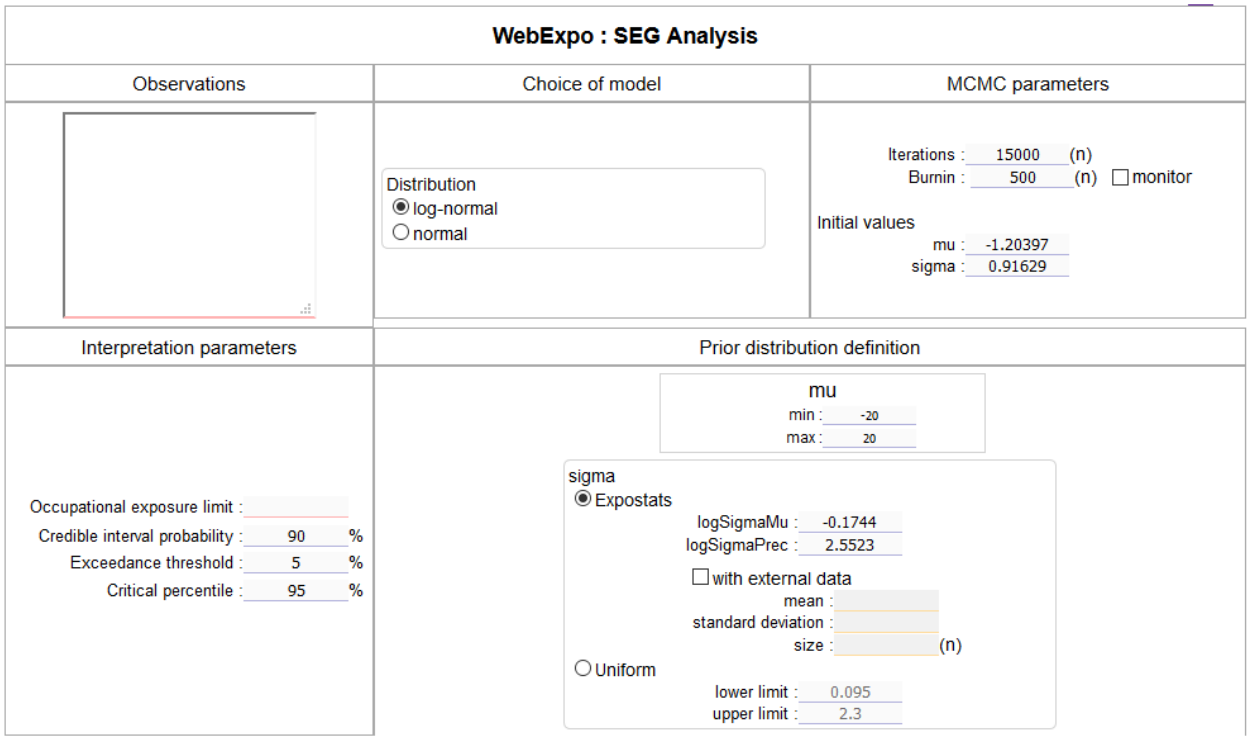


Figure 8. User interface of the JavaScript WebExpo prototype.

5. DISCUSSION

5.1 Overview

Despite the existence of a consensus framework for the analysis of occupational hygiene measurement data, refined up to the last decade and exemplified in the recent European guidelines (CEN, 2018), there is a scarcity of tools available to practitioners to perform the rather complex associated calculations, especially related to data containing non-detects, a frequent occurrence in our field.

The WebExpo project aimed at creating a list of calculations that would provide a comprehensive answer to IH data interpretation based on an ensemble of current guidelines, create algorithms that would implement these calculations within a single methodological framework, and share these algorithms in different languages to facilitate their use for the creation of practical tools. We achieved these goals through establishing the list of calculations based on a review of guidelines and recent literature, and a consultation with a panel of international exposure assessment experts, implementing the calculations using Bayesian statistics, and creating algorithms in statistical and programming languages showcased by two prototype tools.

The resulting WebExpo algorithms allow analysing lognormal or normal data both for estimating a single distribution or within- and between-worker components of variance when repeated measurements are available on some individuals. In each case, several types of prior information can be used, including information from relevant external data and expert judgment. Censored data treatment (left, right or interval) is seamlessly integrated in all calculations, and measurement error can be taken into account in the analysis. Finally, in addition to the more traditional confidence intervals, the nature of Bayesian statistics allowed to express uncertainty in the form of probabilistic statements.

5.2 Choosing between different Bayesian priors

The WebExpo models provide several choices of prior distribution depending on the type of analysis. Mostly, they can be summarized as: the *informedvar* models, where there is little prior information on the mean, but variability is minimally informed based on historical data; the *uninformative* models, where prior distributions are uniform with large (customizable) ranges; and the *riskband* and *past.data* models, where significant information can be present. The traditional Bayesian approach recommends assessing robustness across a range of different priors to widen the interpretation of an analysis (Gelman, 2013), as it will apply to a wider variety of interpretations. For realistic sample sizes in our field (5-10 observations), informative priors will typically have a non-trivial effect on the final exposure estimates (Jones & Burstyn, 2017) compared to uninformed priors. Hygienists analysing industrial hygiene data and trying different priors (uninformative or informed) therefore run the risk of finding themselves in a quandary, with potentially very different results across priors. Of course, such situations would warrant further evaluation efforts. Note that this is a strength rather than a weakness of the Bayesian approach, and it clearly points out that the sample sizes collected to assess workplace exposure are often inadequate. Adequate sample sizes would allow for strong conclusions regardless of the prior information used, most often not the case in this area.

We recommend using the *informedvar* priors as a reasonable default in the WebExpo library as they represent a compromise, with an uninformative prior for the geometric mean, but with a moderately informative prior on variability, based on a population of available workplace variability values, as used by McNally *et al.* and recently advocated by Jones and Burstyn (Jones & Burstyn, 2017; McNally *et al.*, 2014).

5.3 Strengths

To our knowledge, the WebExpo algorithms include the most comprehensive list of calculations considered relevant to the interpretation of industrial hygiene data based on current best practice. While all calculations proposed in our algorithms may not be deemed of interest to all, the open source nature of WebExpo should allow for the creation of applications tailored to any specific need. Moreover, the major contribution of this project being the Bayesian engine and the creation of the posterior samples through MCMC, any additional treatment of these data (outside of the calculations in Tables 2 and 3) is straightforward.

Despite the fact that the treatment of non-detects has long represented a major challenge in the interpretation of IH data, most recent developments (see e.g., Huynh *et al.* and Krishnamoorty and Mathews (Huynh *et al.*, 2016; Krishnamoorthy *et al.*, 2009)) have not been implemented in practical tools. IHSTAT, arguably the most popular data analysis tool in IH, doesn't allow for censored data, although Lavoué recently proposed a tool to implement the regression on order statistics approach within IHSTAT (J. Lavoué, 2013). HYGINIST implements regression on order statistics limited to one censoring point, as does BWSTAT. IHData analyst and ProUCL include several existing procedures, but none based on multiple imputation. WebExpo uses the same Bayesian approach as described in Huynh *et al.* (2014) and implemented in ART (McNally *et al.*, 2014), applied to all models, and extended from only left-censored data (arguably the most common case in industrial hygiene) to interval-censored and right-censored data.

Management of uncertainty is an essential part of risk assessment (Waters *et al.*, 2015). In WebExpo, we leveraged the probabilistic nature of Bayesian statistics to propose, along with the more traditional calculations, a framework where one can estimate the probability that overexposure criteria are met. Inspired by Hewett *et al.* (Paul Hewett *et al.*, 2006), whose proposal we extended to other metrics and types of analyses, we believe that stating exposure data interpretation in the form "there is an XX% chance that our overexposure criterion is met" is more efficient when communicating with managers and workers compared to traditional reporting of statistical complexities.

The WebExpo algorithms and prototypes have a potentially wider application than IH measurement data analysis. Hence, in essence, we created Bayesian engines for the estimation of lognormal and normal parameters of any quantity for which these models may be deemed relevant. Moreover, the between-worker difference analysis model can be used using other grouping units instead of workers, such as establishment, occupation, or contaminated site for environmental pollution data.

Our algorithms are the first to allow taking into account measurement error in IH data analysis. As mentioned in the introduction, only two publications assessed whether analytical variability might impact the evaluation of environmental variability. We do not think that measurement error should always be considered, as it is computationally costly, but, at least, situations when it is deemed important can now be processed rigorously. More importantly perhaps, the possibility to

include measurement error in IH data interpretation without recourse to any simplification should allow revisiting Nicas *et al.* (Nicas *et al.*, 1991) and Grzebyk and Sandino's (Grzebyk & Sandino, 2005) foundation work to obtain a more refined picture of the impact or measurement error on decision-making.

Finally, the WebExpo algorithms are freely available under the open source license Apache 2.0. They can therefore be used without restriction by anyone wanting to create tools or include them in a pre-existing data management system (provided proper acknowledgement to the researchers and funders). We hope this will help their use by institutions and companies working in IH-related fields. While not directly aimed at practitioners, this availability should facilitate the creation of powerful practical data interpretation.

5.4 Limitations

All models set up by the statistics team were not implemented in all platforms. Hence, measurement error is only available in R and C#, and the risk-band prior model is only available in R. These restrictions were necessary given the unforeseen challenge represented by translation of the R code with measurement error. As an illustration, the main MCMC code went from ~500 to ~3 500 lines when measurement error was added. As the common usage in IH data analysis does not consider measurement error, we believe that the majority of current needs would be covered by the models implemented in all platforms in our project. Users wishing to perform more advanced analyses can use the C# or R code. As pure R code is available for everything, extending the current C# or JavaScript capabilities is also possible given appropriate resources.

WebExpo does not include non-parametric statistics, nor does it include functions for the verification of the distributional shape of the sample. Both features were discussed in the expert committee meeting. With the low statistical power associated with distribution-free approaches, it was judged that adding them would not be useful given current practice in terms of sample size (typically <10 to assess a particular situation). In a similar fashion, the practical usefulness of formal hypothesis tests to evaluate normality/lognormality (popular tests in IH include the Shapiro-Wilk and Shapiro-Francia tests (Shapiro & Francia, 1972a, 1972b)) was judged limited in order to decide whether to use or not the lognormal framework in risk assessment. In essence, answering the question "Does this sample come from a lognormal distribution", which is the formal test question, is not directly relevant to the actual question "Does the lognormal model permit drawing useful conclusions". Moreover, the power of such tests is very limited at current sample sizes in IH. As a consequence, we recommend, for the interpretation of workplace exposure data, presuming lognormality as a reasonable default assumption given past empirical evidence, but also using graphical tools such as the Q-Q plot to detect any strong deviation. Notwithstanding the WebExpo team opinion about formal statistical tests, users are free to use or built any procedure to evaluate the data prior to being fed to the lognormal/normal WebExpo calculation algorithms.

Limited "determinants of exposure" analysis tools were initially considered for inclusion in WebExpo. For example, the possibility to analyse the effect of a categorical (ANOVA type analysis) or continuous (linear or smoothed regression) variable, up to a 2 variable model, was examined. Such analyses have indeed become the bread and butter of exposure dataset analyses published in IH literature, and are of increased usefulness as companies/institutions accumulate exposure data in computerized format. They were excluded from the scope of WebExpo, as the experts judged that the level of statistical expertise necessary to adequately

conduct such analyses would imply mastery of a statistical package, and thus render creating a tool a fruitless endeavour. However we would like to note that performing the SEG analysis separately on each category of a nominal variable (e.g., day/night shift data) and combining the MCMC sample would allow performing analyses related to the traditional ANOVA (e.g., estimate difference in exceedance fraction, + credible interval, between any 2 categories). However they are not equivalent, as ANOVA uses the whole dataset as opposed to only strata-specific datasets. In addition, ANOVA assumes a common variance within strata, as opposed to strata-specific variances.

Finally, while we believe the algorithms developed in this project will ultimately benefit IH practitioners, they do not represent an immediate accessible toolbox for data interpretation. Hence, accessible introduction to lognormal statistics, illustrative graphs, and output complexity tailored to the level of expertise of different users were all judged essential elements of a useful IH data interpretation toolbox by both committees. We are in full agreement with this assessment, and would like to underline that the C# and JavaScript prototypes should not be seen as practical IH tools as they contain none of these elements. We believe, however, that we have created a solid and comprehensive computational foundation that should serve as a starting point to create tools that can be tailored to the specific needs of various stakeholders.

5.5 Relationship between Webexpo and the Expostats online data interpretation toolbox

Approximately at the same time when the application for the current project was submitted for funding to IRSST (in 2014), our team at the University of Montreal launched the first iteration of the Expostats toolset[16], a free web-based industrial hygiene data interpretation suite. Initially very limited in terms of user interface and capability, it has evolved into a comprehensive set of tools, now also available for offline use, and has been recently described by Lavoué *et al.* in the Annals of work exposures and health (Jérôme Lavoué *et al.*, 2018). The Expostats tools allow similar calculations as described in this report except they only include one type of prior (the informedvar prior) and do not permit measurement error. The Bayesian models on Expostats are run using JAGS and R scripts and the SHINY[17] application, which serves as an interface between R and users online. The Expostats toolbox runs on servers with limited capability, restraining the number of simultaneous users, and the calculation engine cannot easily be taken up by others for creating their own tools due to licensing issues. In essence, Expostats aimed at providing practitionners with advanced calculation tools in the short term and has been functional for already several years. On the other hand, Webexpo aimed at creating an open source algorithmic foundation for the same set of calculations, enabling institutions or companies to create solutions tailored to their own needs, ultimately leading, in the longer term, to a wider use of state-of-the-art data interpretation practice in our field.

[16] http://www.expostats.ca/site/en/index.html
[17] https://shiny.rstudio.com/

6. CONCLUSION

Quantitative IH data interpretation is just one part of risk analysis in the workplace, and in many cases, decisions can be reached without the need to collect measurements. However, the availability of quantitative exposure data warrants an adequate interpretation. This remains a challenging part of risk assessment in the workplace given high environmental variability, a rather complex statistical analysis framework, and a surprising scarcity of practical tools. The WebExpo project represents a major effort to translate recent computational and theoretical developments towards practice. While the proposed algorithms and prototypes cannot yet be used directly by practitioners, they can be freely used by any institution, individual or corporation as a powerful and rigorous foundation for next generation IH data interpretation tools. In particular, the two prototypes will serve as starting points for the future creation of an IRSST-specific fully-fledged practical data interpretation tool. It remains important to keep in mind that soundness of the conclusions will rely on the suitability of the lognormal/normal model for the situation at hand, on the representativeness and quality of the samples collected, as well as the adequacy of priors.

BIBLIOGRAPHY

Arnold, S. F., Stenzel, M., Drolet, D. and Ramachandran, G. (2016). Using checklists and algorithms to improve qualitative exposure judgment accuracy. *Journal of Occupational and Environmental Hygiene*, *13*(3), 159-168.

Ashley, K. and Bartley, D. L. (2004). Analytical performance criteria. *Journal of Occupational and Environmental Hygiene*, *1*(4), D37-D41.

Banerjee, S., Ramachandran, G., Vadali, M. and Sahmel, J. (2014). Bayesian hierarchical framework for occupational hygiene decision making. *Annals of Occupational Hygiene*, *58*(9), 1079-1093.

Bartley, D. L. (2001). Definition and assessment of sampling and analytical accuracy. *Annals of Occupational Hygiene, 45*(5), *357-364*.

Bartley, D. and Lidén, G. (2008). Measurement uncertainty. *The Annals of Occupational Hygiene*, *52*(6), 413-417.

BOHS-NVvA. (2011). *Testing compliance with occupational exposure limits for airborne substances*. Retrieved from http://www.bohs.org/library/technicalpublications

BOHS Technology Committee Working Group. (1993). *British Occupational Hygiene Society: Technical guide n°. 11 : Sampling strategies for airborne contaminants in the workplace*. *leeds*. Canberra, Australia: Libraries Australia.

Breslin, A. J., Ong, L., Glauberman, H., George, A. C. and Leclare, P. (1967). The accuracy of dust exposure estimates obtained from conventional air sampling. *American Industrial Hygiene Association Journal*, *28*(1), 56-61.

Buringh, E. and Lanting, R. (1991). Exposure variability in the workplace: Its implications for the assessment of compliance. *American Industrial Hygiene Association Journal*, *52*(1), 6-13.

CEN. (1995). *Workplace atmospheres: Guidance for the assessment of exposure by inhalation to chemical agents for comparison with limit values and measurement strategy*. Standard EN 689:1996. Brussels, Belgium: CEN.

CEN. (2018). *Workplace exposure: Measurement of exposure by inhalation to chemical agents: Strategy for testing compliance with occupational exposure limit values*. Standard EN 689:2018. Brussels, Belgium: CEN.

Clerc, F. and Vincent, R. (2014). Assessment of occupational exposure to chemicals by air sampling for comparison with limit values: The influence of sampling strategy. *The Annals of Occupational Hygiene*, *58*(4), 437-449.

Comité européen de normalisation. (1995). *Atmosphères des lieux de travail : conseils pour l'évaluation de l'exposition aux agents chimiques aux fin de comparaison avec des valeurs limites et stratégie de mesurage.* Norme NF EN 689. Bruxelles, Belgique: Communauté européenne.

Drolet, D. and Beauchamp, G. (2013). *Sampling guide for air contaminants in the workplace* (8th ed). (Report n° T-15). Montreal, QC: IRSST

Drolet, D., Goyer, N., Roberge, B., Lavoué, J., Coulombe, M. and Dufresne, A. (2013). *Stratégies de diagnostic de l'exposition des travailleurs aux substances chimiques* (Report n° 665). Montréal, QC: IRSST.

Esmen, N. (1979). Retrospective industrial hygiene surveys. *American Industrial Hygiene Association Journal*, *40*(1), 58-65.

Esmen, N. A. and Hammad, Y. H. (1977). Log-normality of environmental sampling data. *Journal of Environmental Science and Health*, *A12*, 29-41

Espino-Hernandez, G., Gustafson, P. and Burstyn, I. (2011). Bayesian adjustment for measurement error in continuous exposures in an individually matched case-control study. *BMC Medical Research Methodology*, *11*(1), 67.

Flynn, M. R. (2010). Analysis of censored exposure data by constrained maximization of the Shapiro-Wilk W statistic. *The Annals of Occupational Hygiene*, *54*(3), 263-271.

Ganser, G. H. and Hewett, P. (2010). An accurate substitution method for analyzing censored data. *Journal of Occupational and Environmental Hygiene*, *7*(4), 233-244.

Gelman, A. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC Press.

Groth, C., Banerjee, S., Ramachandran, G., Stenzel, M. R., Sandler, D. P., Blair, A., . . . Stewart, P. A. (2017). Bivariate left-censored bayesian model for predicting exposure: Preliminary analysis of worker exposure during the deepwater horizon oil spill. *Annals of Work Exposures and Health*, *61*(1), 76-86.

Grzebyk, M. and Sandino, J. P. (2005). Aspects statistiques et rôle de l'incertitude de mesurage dans l'évaluation de l'exposition professionnelle aux agents chimiques. *Hygiène et sécurité du travail*, *200*, 9-22.

Hawkins, N. C., Norwood, S. K. and Rock, J. C. (1991). *A strategy for occupationnal exposure assessment*. Fairfax, VA: American Industrial Hygiene Association.

Helsel, D. (2005). *Non detects and data analysis: Statistics for censored environmental data*. Hoboken, NJ: John Wiley & Sons.

Helsel, D. (2010). Much ado about next to nothing: Incorporating nondetects in science. *The Annals of Occupational Hygiene*, *54*(3), 257-262.

Helsel, D. R. (2012). *Statistics for censored environmental data using minitab and R* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Hewett, P. (1997). Mean testing: I. Advantages and disadvantages. *Applied Occupational and Environmental Hygiene*, *12*(5), 339-346.

Hewett, P. and Ganser, G. H. (2007). A comparison of several methods for analyzing censored data. *The Annals of Occupational Hygiene*, *51*(7), 611-32.

Hewett, P., Logan, P., Mulhausen, J., Ramachandran, G. and Banerjee, S. (2006). Rating exposure control using Bayesian decision analysis. *Journal of Occupational and Environmental Hygiene*, *3*(10), 568-581.

Hornung, R. and Reed, L. D. (1990). Estimation of average concentration in the presence of nondetectable values. *Applied Occupational and Environmental Hygiene*, *5*(1), 46-51.

Huynh, T., Quick, H., Ramachandran, G., Banerjee, S., Stenzel, M., Sandler, D. P., . . . Stewart, P. A. (2016). A comparison of the β-substitution method and a Bayesian method for analyzing left-censored data. *The Annals of Occupational Hygiene*, *60*(1), 56-73.

Huynh, T., Ramachandran, G., Banerjee, S., Monteiro, J., Stenzel, M., Sandler, D. P., . . . Stewart, P. A. (2014). Comparison of methods for analyzing left-censored occupational exposure data. *The Annals of Occupational Hygiene*, *58*(9), 1126-1142.

Ignacio, J. S. and Bullock, W. H. (2008). *A strategy for assessing and managing occupational exposures* (3rd ed.). Fairfax, VA: AIHA Press.

INRS. (2018). *Interprétation statistique des résultats de mesure.* Paris, France: INRS.

Jahn, S. D., Bullock, C. and Ignacio, J. S. (2015). *A strategy for assessing and managing occupational exposures* (4th ed.). Fairfax, VA: AIHA Press.

Jayjock, M. A., Chaisson, C. F., Franklin, C. A, Arnold, S. and Price, P. S. (2009). Using publicly available information to create exposure and risk-based ranking of chemicals used in the workplace and consumer products. *Journal of Exposure Science & Environmental Epidemiology*, *19*(5), 515-524.

Jones, R. M. and Burstyn, I. (2017). Bayesian analysis of occupational exposure data with conjugate priors. *Annals of Work Exposures and Health*, *61*(5), 504-514.

Kerr, G. W. (1962). Use of statistical methodology in environmental monitoring. *American Industrial Hygiene Association Journal*, *23*(1), 75-82.

Krishnamoorthy, K., Mallick, A. and Mathew, T. (2009). Model-based imputation approach for data analysis in the presence of non-detects. *The Annals of Occupational Hygiene*, *53*(3), 249-263.

Kromhout, H., Symanski, E. and Rappaport, S. M. (1993). A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. *The Annals of Occupational Hygiene*, *37*(3), 253-270.

Kumagai, S. and Matsunaga, I. (1995). Changes in the distribution of short-term exposure concentration with different averaging times. *American Industrial Hygiene Association Journal*, *56*(1), 24-31.

Lavoué, J. (2013). Dealing with non-detects in occupational hygiene datasets. *Exposure,* December 2013, 13-16.

Lavoue, J., Friesen, M. C. and Burstyn, I. (2013). Workplace measurements by the US occupational safety and health administration since 1979: Descriptive analysis and potential uses for exposure assessment. *Annals of Occupational Hygiene*, *57*(1), 77-97.

Lavoué, J., Joseph, L., Knott, P., Davies, H., Labrèche, F., Clerc, F., . . . Kirkham, T. (2018). Expostats: A Bayesian toolkit to aid the interpretation of occupational exposure measurements. *Annals of Work Exposures and Health*. Retrived from http://doi.org/10.1093/annweh/wxy100

Leidel, N. A. and Busch, K. A. (2000). Statistical design and data analyis. In R. L. Harris (Ed.), *Patty's industrial hygiene* (5th ed., p. 2387-2514). New York, NY: John Wiley & Sons.

Leidel, N. A., Busch, K. A. and Lynch, C. F. (1977). *NIOSH Occupational exposure sampling strategy manual*. Cincinnati, OH: US Department of Health, Education, and Welfare.

Leidel, N., Busch, K. and Crouse, W. E. (1975). NIOSH *Technical information : Exposure measurement action level and occupational environmental variability*: NIOSH 76-131. Cincinnati, OH: NIOSH. Retrived from https://www.cdc.gov/niosh/docs/76-131/pdfs/76-131.pdf

Logan, P., Ramachandran, G., Mulhausen, J. and Hewett, P. (2009). Occupational exposure decisions: Can limited data interpretation training help improve accuracy? *The Annals of Occupational Hygiene*, *53*(4), 311-324.

Logan, P. W., Ramachandran, G., Mulhausen, J. R., Banerjee, S. and Hewett, P. (2011). Desktop study of occupational exposure judgments: Do education and experience influence accuracy? *Journal of Occupational and Environmental Hygiene*, *8*(12), 746-758.

Lyles, R. H. and Kupper, L. L. (1996). On strategies for comparing occupational exposure data to limits. *American Industrial Hygiene Association Journal*, *57*(1), 6-15.

Lyles, R. H., Kupper, L. L. and Rappaport, S. M. (1997a). A lognormal distribution-based exposure assessment method for unbalanced data. *Annals of Occupational Hygiene*, *41*(1), 63-76.

Lyles, R. H., Kupper, L. L. and Rappaport, S. M. (1997b). Assessing regulatory compliance of occupational exposures via the balanced one-way random effects ANOVA model. *Journal of Agricultural, Biological, and Environmental Statistics*, *2*(1), 64-86.

Martin Remy, A. and Wild, P. (2017). Bivariate left-censored measurements in biomonitoring: A Bayesian model for the determination of biological limit values based on occupational exposure limits. *Annals of Work Exposures and Health*, *61*(5), 515–527.

Mcbride, S., Williams, R. and Creason, J. (2007). Bayesian hierarchical modeling of personal exposure to particulate matter. *Atmospheric Environment*, *41*(29), 6143-6155.

McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton, FL: CRC Press.

McNally, K., Warren, N., Fransman, W., Entink, R. K., Schinkel, J., van Tongeren, M., . . . Tielemans, E. (2014). Advanced REACH Tool: A Bayesian model for occupational exposure assessment. *The Annals of Occupational Hygiene*, *58*(5), 551-565.

Morton, J., Cotton, R., Cocker, J. and Warren, N. D. (2010). Trends in blood lead levels in UK workers, 1995-2007. *Occupational and Environmental Medicine*, *67*(9), 590-595.

Mulhausen, J. R. and Diamano, J. (1998). *A strategy for assessing and managing occupational exposures* (2nd ed.). Fairfax, VA: AIHA Press.

Nicas, M., Simmons, B. P. and Spear, R. C. (1991). Environmental versus analytical variability in exposure measurements. *American Industrial Hygiene Association Journal*, *52*(12), 553–557.

Ogden, T. L. (2010). Handling results below the level of detection. *The Annals of Occupational Hygiene*, *54*(3), 255-256.

Ogden, T. and Lavoué, J. (2012). Testing compliance with occupational exposure limits: Development of the British-Dutch guidance. *Journal of Occupational and Environmental Hygiene*, *9*(4), D63-70.

Oldham, P. D. (1953). The nature of the variability of dust concentrations at the coal face. *British Journal of Industrial Medicine*, *10*(4), 227-234.

Pesch, B., Kendzia, B., Hauptmann, K., Van Gelder, R., Stamm, R., Hahn, J.-U., . . . Brüning, T. (2015). Airborne exposure to inhalable hexavalent chromium in welders and other occupations: Estimates from the German MEGA database. *International Journal of Hygiene and Environmental Health*, *218*(5), 500-506.

Pilote, L., Joseph, L., Bélisle, P., Robinson, K., Van Lente, F. and Tager, I. B. (2000). Iron stores and coronary artery disease: A clinical application of a method to incorporate measurement error of the exposure in a logistic regression model. *Journal of Clinical Epidemiology*, *53*(8), 809-816.

Quick, H., Huynh, T. and Ramachandran, G. (2017). A method for constructing informative priors for Bayesian modeling of occupational hygiene data. *Annals of Work Exposures and Health*, *61*(1), 67-75.

R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Ramachandran, G. (2008). Toward better exposure assessment strategies: The new NIOSH initiative. *The Annals of Occupational Hygiene*, *52*(5), 297-301.

Ramachandran, G. and Vincent, J. H. (1999). A Bayesian approach to retrospective exposure assessment. *Applied Occupational and Environmental Hygiene*, *14*(8), 547–557.

Rappaport, S. M. (1984). The rules of the game: An analysis of OSHA's enforcement strategy. *American Journal of Industrial Medicine*, *6*(4), 291-303.

Rappaport, S. M. (1991). Assessment of long-term exposures to toxic substances in air. *The Annals of Occupational Hygiene*, *35*(1), 61-121. Retrived from http://doi.org/10.1093/annhyg/35.6.674

Rappaport, S. M. (2000). Interpreting levels of exposures to chemical agents. In R. L. Harris (Ed.), *Patty's industrial hygiene* (5th ed., p. 679-745). New York, NY: John Wiley & Sons.

Rappaport, S. M., Kromhout, H. and Symanski, E. (1993). Variation of exposure between workers in homogeneous exposure groups. *American Industrial Hygiene Association Journal*, *54*(11), 654-662.

Rappaport, S. M., Lyles, R. H. and Kupper, L. L. (1995). An exposure-assessment strategy accounting for within- and between-worker sources of variability. *Annals of Occupational Hygiene*, *39*(4), 469-495.

République française. (2009). Arrêté du 15 décembre 2009 relatif aux contrôles techniques des valeurs limites d'exposition professionnelle sur les lieux de travail et aux conditions d'accréditation des organismes chargés des contrôles. *Journal officiel de la République française*, Texte 35 sur 156.

Roach, S. A. (1966). A more rational basis for air sampling progammes. *American Industrial Hygiene Association Journal*, *27*(1), 1-12.

Roach, S. A. (1977). A most rational basis for air sampling programmes. *The Annals of Occupational Hygiene*, *20*(1), 65-84.

Sarazin, P., Labrèche, F., Lesage, J. and Lavoué, J. (2018). *Étude comparative des banques de données de mesures d'exposition IMIS (OSHA) et LIMS* (Rapport n° R-1032). Montréal, QC: IRSST.

Shapiro, S. S. and Francia, R. (1972a). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, *67*(337), 215-216.

Shapiro, S. S. and Francia, R. S. (1972b). An approximate analysis of variance test for normality. *Journal of American Statistical Association, 67*(32), 215-216.

Sottas, P.-E., Lavoué, J., Bruzzi, R., Vernez, D., Charrière, N. and Droz, P.-O. (2009). An empirical hierarchical Bayesian unification of occupational exposure assessment methods. *Statistics in Medicine*, *28*(1), 75-93.

Tornero-Velez, R., Symanski, E., Kromhout, H., Yu, R. C. and Rappaport, S. M. (1997). Compliance versus risk in assessing occupational exposures. *Risk Analysis*, *17*(3), 279-292.

Waters, M., McKernan, L., Maier, A., Jayjock, M., Schaeffer, V. and Brosseau, L. (2015). Exposure estimation and interpretation of occupational risk: Enhanced information for the occupational risk manager. *Journal of Occupational and Environmental Hygiene*, *12*(S1), S99-S111.

Wild, P., Hordan, R., Leplay, A. and Vincent, R. (1996). Confidence intervals for probabilities of exceeding threshold limits with censored log-normal data. *Environmetrics*, *7*(3), 247-259.

Zhang, Y. F., Banerjee, S., Yang, R., Lungu, C. and Ramachandran, G. (2009). Bayesian modeling of exposure and airflow using two-zone models. *The Annals of Occupational Hygiene*, *53*(4), 409-424.

ANNEXE A:
MEETING NOTES FROM THE INTERNATIONAL EXPERT MEETING

WebExpo meeting
February 18-19, 2016
Montréal
Final meeting notes

CONTENT

1. INTRODUCTION

This document summaries the discussions during the WebExpo meeting, held February 18-19, 2016, in Montréal. It was drafted by Jérôme Lavoué and commented/revised by the meeting attendees. Very little background information is provided since it is available in the scientific protocol and in the meeting preparatory documents. This summary is written so as to present the general impressions/ideas collected during the meeting rather than specific arguments by particular persons.

Table A1: list of attendees at the expert meeting

Name	Affiliation	Domain of expertise	Title
Jérôme Lavoué	Université de Montréal	Exposure science	PI
Lawrence Joseph	McGill University	Statistics	Co-PI
Simon Aubin	IRSST	IH metrology	Collaborator
France Labrèche	IRSST	Epidemiology	Collaborator
Tracy Kirkham	University of Toronto	Industrial hygiene	Collaborator
Gautier Mater	INRS	Industrial hygiene	Collaborator
Frédéric Clerc	INRS	Statistics	Collaborator
Patrick Belisle	McGill University	Statistical programming	Research officer
Dunia Ouedraogo	Université de Montréal	Exposure science	Ph.D student
Martie van Tongeren	IOM	Exposure science	Expert
Martine Chouvet	ITGA	Industrial hygiene	Expert
Paul Bozek	University of Toronto	Industrial hygiene	Expert
Hugh Davies	University of British Columbia	Industrial hygiene	Expert
Michel Gérin	Université de Montréal	Industrial hygiene	Expert

2. LIST OF CORE CALCULATIONS – FUNCTIONALITIES INCLUDED IN WebExpo

2.1 Group assessment (e.g., data from a similar exposure group)

There was a large consensus that the 3 proposed risk metrics: exceedance (needs specification of the OEL), 95th percentile, and arithmetic mean should be included in WebExpo.

2.2 Between- and within-worker variability model

The group agreed that the following metrics were relevant for the between- and within-worker analysis:

- Probability that a random worker has his own 95th percentile above the OEL
- Probability that a random worker has his own arithmetic mean above the OEL
- Between-worker variability
- Within-worker variability (one mean value for all workers)
- Within-worker correlation coefficient
- Rappaport's ratio

Other proposals to reflect differences between workers included reporting an interval or standard deviation for the risk metrics of workers, e.g., average exceedance was 15%, with a standard deviation of 5% across workers. A tweaking of the Rappaport's ratio was also proposed, using less extreme percentiles of the between-worker distribution, e.g., 10% and 90% rather than 2.5% and 97.5%, in order to be more easily interpretable for smaller groups of workers (i.e., what is the 2.5th percentile of a population of 10 workers?).

In terms of evaluating whether a group exposure is homogenous or not, the usefulness of the ANOVA test was criticised strongly because: its real meaning is not understood by most, and because it does not provide an answer to the actual question of interest "Are differences between workers sufficient to have an impact on the final diagnosis?" On the one hand, small sample size won't allow detecting important differences; on the other hand, rejection of the null hypothesis can occur even if the differences are very small. The group consensus is that providing estimates of the amplitude of between-worker differences is preferable.

The group also agreed that it is useful to provide individual exposure estimates for each sampled worker, although these estimates are quite uncertain. Ethical considerations related to "pointing fingers" should nevertheless be taken into account when reporting/communicating the results.

It was underlined that although it would be interesting to estimate worker-specific within-worker variability, it was not realistic to do so with usual sample sizes. There was an interest in simulation projects exploring the impact of failing to model such differences in diagnosis.

2.3 Other functionalities

The group agreed that it would be interesting to include the possibility to model the influence of one categorical variable. It should be flexible and allow input of data through EXCEL files with several candidates that could be analysed separately by a simple selection by the user.

For the potential addition of a continuous variable, the group expressed interest (sample duration, temporal trend) but acknowledged the added technical difficulties (adequacy of a simple slope, interpretation of the slope). The biggest interest was in temporal trends.

The group did not express a huge interest in the possibility to analyse interval censored data (rather than only left censored data), or to evaluate serial correlation.

Some institutions recommend flagging situations with an elevated GSD to identify abnormal conditions. There was not a consensus in the group about the interest of this procedure, or to which values would represent “acceptable” GSDs. An elevated GSD may only reflect very variable exposure conditions (e.g., firefighting).

The possibility to include sampling error was raised, as it is easy to model in the Bayesian framework and uncertainty around a full-shift value might be important when only a part of the shift was actually monitored. As available studies (2 papers) showed that measurement error in IH would matter in very few situations compared to environmental variability, the group agreed that it was not necessarily useful to model the measurement error across the board, but better to offer the possibility to include it in restricted conditions.

3. DISCUSSION AROUND BAYESIAN PRIORS

It was mentioned that most users won’t have the prerequisite knowledge to evaluate the adequacy/value of different proposed priors. Generic priors might be applicable in several situations (e.g., GSD based on the Kromhout *et al.* database), but will bring very little information. On the other hand, informative specific priors (e.g., BDA style) can bring more information, but it might be very costly to make sure they are accurate. In addition, it will not be possible to know if they really are accurate, given that the actual data will probably be insufficient to check that.

We are in a situation (small sample size) where it is not really possible to assess whether the various priors are good, and where the final results will rely quite heavily on the priors. As a result, if they do not agree, the only valid conclusion will be that the data is not good enough and that more samples are required.

About priors that resemble another dataset (i.e., provide a GM, GSD and a virtual sample size), it is remarked that users would typically be lost when they have to select the virtual sample size. As an alternative, it would be possible to have a slider showing what happens for different values of these parameters.

It is also remarked that rather than only having a parameter reflecting virtual sample size, it would be better to have a parameter reflecting “closeness” of the prior values to the situation at hand (e.g., a slider from “relevant to my situation” to “little relevant”).

The statisticians in the group underlined that data typically collected in IH does not seem sufficient to adequately characterize exposure in many situations. It is a consensus in the group that the tools should then focus on illustrating the important uncertainty in exposure estimates. One possibility would involve taking the focus away from point estimates, to reflect more the range of likely values.

The group agrees that there is interest in including several different priors in the tool, but that it is a challenge to present them and provide support for the interpretation of the results (inappropriate priors can bias results). It is suggested that studies be made to look at how the results from different priors compared and to hopefully serve as a basis for guidance.

Regarding the BDA/AIHA prior bands, the group deemed them difficult to use to elicit a prior from users ("uneven broad categories", based on unfamiliar parameters, P95). It was judged that users are better at estimating central tendencies. This scheme might be useful to present results, but seems less interesting as a way to elicit good priors.

4. DISCUSSION AROUND SMALL DATASETS

INRS states that an absolute minimal sample size of 3 is recommended in their guidelines. For samples smaller than 6, there is no calculation of distributional parameters; rather, the maximum of the series of value is compared to a fraction of the OEL. This approach relies on an *a priori* value for the GSD, not estimated from the data.

It is remarked that for small sample sizes, confidence intervals based on frequentist calculation might not be reliable, which supports the idea of an alternative approach for these situations. It is underlined, however, that Bayesian credible intervals are not affected by this problem since uncertainty about the mean and standard deviation are fully taken into account in the priors. An approach based on a Bayesian framework would therefore not require treating very small datasets differently.

5. DISCUSSION AROUND DISTRIBUTION FREE APPROACHES – ASSUMPTIONS ABOUT THE LOGNORMAL MODEL

The consensus around this issue was that:

- Hypothesis tests such as the Shapiro-Wilk test are not useful to evaluate whether the underlying distribution is lognormal:
  - They do not tell "how far" we are from the lognormal, or whether this departure has any consequence on the interpretation;
  - The distributional shape cannot be examined at "current" sample sizes (5-10);
- For sample sizes below 20-30, the Q-Q plot is not very useful either;
- Uncertainty bands on the Q-Q plot do not help as they will be influenced by extreme points to the extent that a very extreme deviation would be required for a point to fall outside.

In conclusion, formal hypothesis tests will probably not be included in the WebExpo toolset. According to the statisticians in the group, it would not be useful either to try to study distributional shape using QQ plot below 30 points. Below this threshold, simpler plots can help

identify outliers. The choice of the distributional model would therefore be based on *a priori* decision (e.g., lognormal for airborne chemical exposures, normal for noise). QQ plot will probably be included, with warnings that they may not be very helpful at small values of n.

Distribution free approaches were briefly discussed. It was mentioned that they would unlikely be useful at current sample sizes because of the significant loss of power. A simple sequential graph with different symbols for different workers would provide a nice summary and would help identify extreme values.

6. DISCUSSIONS AROUND RISK COMMUNICATION

Regarding the notion of sending different messages to different audiences, the group strongly agreed in favor of this proposal. Colleagues from INRS shared their experience of creating a quiz to select the complexity of the message based on answers. In retrospect they would not recommend the same approach for WebExpo. The group agreed that an approach of the kind "click here for a more detailed interpretation" would work well. This process should make sure that advanced users should not repeatedly have to make this selection. There was also a consensus that the tool should encourage people to want to get the most complex answer. The notion of a quiz was popular not for auto-selecting the complexity of the data interpretation results, but as a part of educational material.

In terms of graphics to illustrate data interpretation results, illustration of exceedance with calendars with greyed out days was popular. Other suggestions included the density curve of the estimated distribution, even a cloud of lognormal density curves to reflect uncertainty. The type of graph should depend on the degree of understanding of the users (e.g., a density curve is not straightforward for many people). The group agreed that a "lognormal simulator" component would be a good thing for educational purposes.

In terms of the diffusion of the algorithms developed, it was underlined that although a wiki page might work well for documents (e.g., examples, guidelines, educational material), it wasn't likely to succeed for computer codes. Moreover, maintenance and updates of a wiki system require a regular source of funding (difficult to ensure here). The various algorithms will use only open access libraries, so that no issue will arise with their distribution and usage. Java versions of the algorithms will be produced, in part because colleagues from INRS plan to use them for the next iteration of their own data interpretation tool, ALTREX2.

7. REFERENCES

1. Kromhout H, Symanski E, Rappaport SM. A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. Ann Occup Hyg 1993; 37:253–70. http://www.ncbi.nlm.nih.gov/pubmed/8346874 (accessed Oct. 9, 2014).

2. Banerjee S, Ramachandran G, Vadali M, *et al.* Bayesian Hierarchical Framework for Occupational Hygiene Decision Making. Ann Occup Hyg 2014; 58:1079-1093. doi:10.1093/annhyg/meu060.

3. McNally K, Warren N, Fransman W, *et al.* Advanced REACH Tool: a Bayesian model for occupational exposure assessment. Ann Occup Hyg 2014; 58:551–65. doi:10.1093/annhyg/meu017.

4. Arnold SF, Stenzel M, Drolet D, *et al.* Using checklists and algorithms to improve qualitative exposure judgment accuracy. J Occup Environ Hyg 2016; 13:159–68. doi:10.1080/15459624.2015.1053892.

5. Logan P, Ramachandran G, Mulhausen J, *et al.* Occupational exposure decisions: can limited data interpretation training help improve accuracy? Ann Occup Hyg 2009; 53:311–24. doi:10.1093/annhyg/mep011.

6. Vadali M, Ramachandran G, Mulhausen JR, *et al.* Effect of training on exposure judgment accuracy of industrial hygienists. J Occup Environ Hyg 2012; 9:242–56. doi:10.1080/15459624.2012.666470.

ANNEXE B: TECHNICAL DOCUMENTATION OF THE BAYESIAN MODELS

WebExpo: The R algorithms

August 2, 2018

1 Introduction

This document addresses sampling from the posterior distributions from several models in Industrial Hygiene. In each model, the prior distribution $f(\theta)$ for the hyperparameters $\theta = (\mu, \sigma)$ is relatively simple.

1.1 Generating a sample from posterior distribution via Markov Chain Monte Carlo (MCMC)

From Bayes theorem, the posterior distribution $f(\theta|x)$ for θ given data x is proportional to

$$f(\theta|x) \propto f(\theta) \times f(x|\theta) \ ,$$

where $f(\theta)$ is the prior distribution for θ and $f(x|\theta)$ is the likelihood function.

In most situations encountered in this work, the posterior for θ does not have an analytic solution but we can use Markov Chain Monte Carlo simulation to draw a sample from it. When θ consists of a series of parameters, say $\theta = (\theta_1, \theta_2, \ldots, \theta_q)$, if the full conditional posterior distribution for θ_i can be written as a function of other components, that is, if we can write $f(\theta_i|\theta_{-i}, x)$, for $i = 1, 2, \ldots, q$ — where $\theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_q)$ — then the MCMC algorithm is as follows: sample θ_i from the above distribution for $i = 1, 2, \ldots, q$, collect the sampled values and repeat a large number of times; in the long run, the sample collected along these lines converges to a sample from the posterior distribution $f(\theta|x)$.

Sections 2–5 present all the models in absence of measurement error while Section 6 presents the modifications to bring to each of them in the presence of measurement error.

2 Uninformative model [SEG.uninformative]

The joint prior distribution for μ and σ in the uninformative model can be constructed from

$$\begin{aligned} \mu &\sim U(\mu_0, \mu_1) \\ \tau = \frac{1}{\sigma^2} &\sim \text{Gamma}(\alpha, \beta) \end{aligned} \tag{1}$$

with $\mu_0 = -1000, \mu_1 = 1000$ and $\alpha = \beta = 0.001$. The likelihood is

$$Y_i \sim N(\mu, \sigma^2)$$

for $i = 1, 2, \ldots, N$, where the Y_i are independenty (and identically) distributed.

The joint posterior distribution for (μ, σ) is thus

$$\begin{aligned} f(\mu, \sigma|y) &\propto \frac{1}{\sigma^N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \mu)^2\right\} \frac{1}{\sigma^{2\alpha+1}} \exp -\frac{\beta}{\sigma^2} I_\mu(\mu_0, \mu_1) \\ &= \frac{1}{\sigma^{N+2\alpha+1}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \mu)^2\right\} \exp -\frac{\beta}{\sigma^2} I_\mu(\mu_0, \mu_1) \end{aligned}$$

where $I_\theta(a,b)$ is the indicator function, that is,

$$I_\theta(a,b) = \begin{cases} 1 & \text{when } a \le \theta \le b \text{ and} \\ 0 & \text{elsewhere.} \end{cases}$$

From the above, we easily derive the full conditional posterior density for μ to get

$$f(\mu|\sigma, y) \sim N(\bar{y}, \sigma^2/N) \tag{2}$$

truncated to (μ_0, μ_1), while the full conditional posterior density for σ is

$$f(\sigma|\mu, y) \quad \propto \quad \frac{1}{\sigma^{N+2\alpha+1}} \exp -\frac{1}{\sigma^2}\left\{\beta + \frac{1}{2}\sum (y_i - \mu)^2\right\},$$

that is, $\tau = \sigma^{-2} \sim \text{Gamma}(\alpha + N/2, \beta + \frac{1}{2}\sum(y_i - \mu)^2)$ from (B.2).

NOTE: After discussion, the gamma prior distribution for τ in (1) was dropped from the package in favor of the uniform prior on σ introduced in next section.

2.1 Alternative posterior when the prior distribution for σ is (improper) uniform

An alternative to the above model is to use a uniform prior distribution on σ rather than the Gamma prior used in (1), that is, to use

$$\sigma \sim U(\sigma_0, \sigma_1)$$

where the range may be left unspecified, that is, with $\sigma_0 = 0$ and $\sigma_1 = \infty$, in which case σ's prior is said to be *improper*, that is, its density does not integrate to 1; this is not a problem in theory but in practice it may leave a high probability for large values for σ — especially when sample size is small — which do not make sense in practice. Hence we encourage the use of a finite upper bound $\sigma_1 < \infty$ based on the scale of the measurements taken.

The conditional posterior density for σ is then

$$f(\sigma|\mu, y) \propto \frac{1}{\sigma^N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \mu)^2\right\}, \tag{3}$$

that is,

$$\begin{aligned} \tau = \sigma^{-2} \quad &\sim \quad \text{Gamma}((N-1)/2, b) \text{ from } (B.2) \\ &\text{where } b = \frac{1}{2}\sum_{i=1}^{N}(y_i - \mu)^2 \text{ when } N > 1. \end{aligned}$$

When $N \le 1$, σ values can be sampled from the inverse cumultative density function (icdf) method presented in Appendix A, with $a = N = 1$ and b defined as above.

3 Kromhout model [SEG.informedvar]

The joint prior distribution for the Kromhout model is given by

$$\begin{aligned} \mu \quad &\sim \quad U(\mu_0, \mu_1) \\ \log(\sigma) \quad &\sim \quad N(\mu^*, \sigma^{*2}) \end{aligned} \tag{4}$$

and the likelihood is

$$Y_i \sim N(\mu, \sigma^2)$$

where the Y_i's, $i = 1, 2, \ldots, N$ are independently distributed and can be left-, right- or interval-censored. The hyperparameter values are $\mu_0 = -101.38161, \mu_1 = 98.61839$, $\mu^* = -0.1744$ and $\sigma^{*-2} = 2.5523$.

The joint posterior for (μ, σ) is hence given by

$$f(\mu, \sigma|y) \propto \frac{1}{\sigma^N} \exp\left\{-\frac{1}{2\sigma^2}\sum(y_i-\mu)^2\right\}\frac{1}{\sigma}\exp\left\{-\frac{(\log(\sigma)-\mu^*)^2}{2\sigma^{*2}}\right\}I_\mu(\mu_0,\mu_1) \tag{5}$$

The full conditional posterior density for μ is thus given by

$$\begin{aligned} f(\mu|\sigma, y) &\propto \exp\left\{-\frac{1}{2\sigma^2}\left(\sum y_i^2 - 2\mu\sum y_i + N\mu^2\right)\right\}I_\mu(\mu_0,\mu_1) \\ &\propto \exp\left\{-\frac{N}{2\sigma^2}\left(\mu^2 - 2\mu\bar{y}\right)\right\}I_\mu(\mu_0,\mu_1)\ , \end{aligned} \tag{6}$$

that is, $\mu \sim N(\bar{y}, \sigma^2/N)$ truncated to the interval (μ_0, μ_1) and the full conditional posterior density for σ is proportional to

$$f(\sigma|\mu, y) \propto \frac{1}{\sigma^{N+1}} \exp\left\{-\frac{1}{2\sigma^2}\sum(y_i-\mu)^2\right\}\exp\left\{-\frac{(\log(\sigma)-\mu^*)^2}{2\sigma^{*2}}\right\} \tag{7}$$

Generating MCMC values for μ from its full conditional posterior density (6) is straightforward, while σ values will be sampled from its full conditional posterior density (7) through the inverse cumulative density function sketched in Appendix A, with

$$\begin{aligned} a &= N, \\ b &= \frac{1}{2}\left(\sum y_i^2 - 2\mu\sum y_i + N\mu^2\right), \\ \tilde{\mu} &= \mu^* \text{ and} \\ \tilde{\sigma}^2 &= \sigma^{*2}\ . \end{aligned}$$

If there are any right-censored values y_i, that is, values specified as $y_i < z_i$ for some z_i's, then at each loop in the MCMC process, corresponding y_i values are sampled from $N(\mu, \sigma^2)$ on the interval $)-\infty, z_i)$. Similar sampling is also performed for left- and interval-censored y_i values.

3.1 Two-Level Kromhout model

The Two-Level Kromhout model is the same as Kromhout model discussed above but applied to two groups, that is, it is a model with the following prior distributions

$$\begin{aligned} \mu_j &\sim U(\mu_0, \mu_1) \\ \log(\sigma_j) &\sim N(\mu^*, \sigma^{*2}) \end{aligned}$$

independently for groups $j = 1, 2$ and the likelihood is given by

$$Y_{ji} \sim N(\mu_j, \sigma_j^2)$$

for $i = 1, 2, \ldots, N_j, j = 1, 2$. The hyperparameter values are $\mu^* = -0.1744$ and $\sigma^{*-2} = 2.5523$ with limits for μ slightly different from the one-group model, $\mu_0 = -100$ and $\mu_1 = 100$.

3.2 Use of past data

One might want to include past data — available through sample size n, observed mean $\bar{p}$ and standard deviation s_p — in the analysis. The likelihood of past data $\mathbf{p}$ — measured without error

— is given by

$$\begin{aligned} f(p|\mu,\sigma) &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(p_i-\mu)^2\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(p_i-\bar{p}+\bar{p}-\mu)^2\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[(p_i-\bar{p})^2+2(p_i-\bar{p})(\bar{p}-\mu)+(\bar{p}-\mu)^2]\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(p_i-\bar{p})^2\right\}\exp\left\{-\frac{n}{2\sigma^2}(\bar{p}-\mu)^2\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{(n-1)s_p^2}{2\sigma^2}\right\}\exp\left\{-\frac{n}{2\sigma^2}(\bar{p}-\mu)^2\right\} \quad . \end{aligned} \tag{8}$$

The joint posterior for (μ,σ) is hence given by the product of (5) and the above likelihood of past data, that is,

$$\begin{aligned} f(\mu,\sigma|y,p) &\propto \frac{1}{\sigma^{N+n+1}}\exp\left\{-\frac{1}{2\sigma^2}\sum(y_i-\mu)^2\right\} \\ &\times \exp\left\{-\frac{(n-1)s_p^2}{2\sigma^2}\right\}\exp\left\{-\frac{n}{2\sigma^2}(\bar{p}-\mu)^2\right\} \\ &\times \exp\left\{-\frac{(\log(\sigma)-\mu^*)^2}{2\sigma^{*2}}\right\}I_\mu(\mu_0,\mu_1) \quad . \end{aligned}$$

The full conditional posterior density for μ is thus given by

$$\begin{aligned} f(\mu|\sigma,y,p) &\propto \exp\left\{-\frac{1}{2\sigma^2}\left[\sum(y_i-\mu)^2+n(\bar{p}-\mu)^2\right]\right\} \\ &\propto \exp\left\{-\frac{1}{2\sigma^2}\left(N\mu^2-2\mu\sum y_i+n\mu^2-2\mu n\bar{p}\right)\right\} \\ &\propto \exp\left\{-\frac{1}{2\sigma^2}\left(\mu^2(N+n)-2\mu(N\bar{y}+n\bar{p})\right)\right\}I_\mu(\mu_0,\mu_1) \\ \Longrightarrow \mu|\sigma,y,p &\sim \mathrm{N}\left(\frac{N\bar{y}+n\bar{p}}{N+n},\frac{\sigma^2}{N+n}\right)I_\mu(\mu_0,\mu_1) \end{aligned}$$

while the full conditional posterior distribution for σ is

$$f(\sigma|\mu,y,p) \propto \frac{1}{\sigma^{N+n+1}}\exp\left\{-\frac{1}{2\sigma^2}\left(\sum(y_i-\mu)^2+(n-1)s_p^2+n(\bar{p}-\mu)^2\right)\right\}\exp\left\{-\frac{(\log(\sigma)-\mu^*)^2}{2\sigma^{*2}}\right\} .$$

Values for σ can be sampled from its full conditional posterior density through the inverse cumulative density function sketched in Appendix A, with

$$\begin{aligned} a &= N+n, \\ b &= \frac{1}{2}\left(\sum y_i^2-2\mu\sum y_i+N\mu^2+(n-1)s_p^2+n(\bar{p}-\mu)^2\right), \\ \tilde{\mu} &= \mu^* \text{ and} \\ \tilde{\sigma}^2 &= \sigma^{*2} \quad . \end{aligned}$$

Originally, the algorithm including past data was based on a uniform and improper distribution for σ which is not used anymore; the rationale for this deprecated version of the algorithm is relegated to Appendix C.5.

3.2.1 Limitations / warnings

If it is thought that the past data were measured with error, they sould NOT be used (indeed, the above section assumed that the past data was measuread without measurement error).

If the measurement error (in past data) was proportional to true (unmeasured) values — that is, measurement error would be modeled through a coefficient of variation — they sould DEFINITELY not be used (the assumptions on which the algorithm is based seem to be violated in a unfixable fashion).

If measurement error (in past data, again) was constant and relatively small when compared to σ, they could still be used, but with some caution. Indeed, the above calculations intrinsically assume that $(n-1)s_p^2/\sigma^2 \sim \chi^2_{n-1}$, which is NOT the case when past values are measured with error. If the measurement error is small, then we may not be very far from that distribution and the algorithm and past data still provide useful results.

The only way that past data obtained with measurement error could be used is if we have access to the complete list of observed values $p_1, p_2, \ldots, p_n$ rather than the usual summary statistics $(\bar{p}, s_p^2)$. In this case, if we additionally assume that the measurement error in past data was of same nature and size as in our actual data $(y_1, y_2, \ldots, y_N)$, then one could simply include past data as (additional) new data. The case when the measurement error is of different nature and/or size than the measurement error in the current data is beyond the scope of this program.

If the outcome of interest follows a log-normal distribution (rather than a normal distribution), then the mean and standard deviation of past data must have been calculated on the log values as well in order to be usable.

4 McNally model [Between.worker]

McNally's model is a hierarchical model with overall mean μ, having prior distribution

$$\mu \sim U(\mu_0^*, \mu_1^*) \, .$$

The model includes random worker effects $\mu_j, j = 1, \ldots, M$ with independent and identical prior distributions

$$\mu_j \sim N(0, \sigma_B^2) \, .$$

The parameter σ_B^2 is the between-worker variance, with prior distribution

$$\log(\sigma_B) \sim N(\mu_B^*, \sigma_B^{*2}) \tag{9}$$

where $\mu_B^* = -0.8786$ and $\sigma_B^{-2*} = 1.634$. The likelihood is given by

$$Y_i \sim N(\mu + \mu_{w_i}, \sigma_W^2)$$

where w_i is the worker index of observation $i, i = 1, \ldots, N$ and σ_W^2 is the within-subject variance, with prior distribution

$$\log(\sigma_W) \sim N(\mu^*, \sigma^{2*}) \tag{10}$$

where $\mu^* = -0.4106$ and $\sigma^{-2*} = 1.9002$.

The joint posterior density is proportional to

$$\begin{aligned} f(\mu, \mu_1, \ldots, \mu_M, \sigma_W^2, \sigma_B^2 | y) \quad \propto \quad & \frac{1}{\sigma_W^N} \exp\left\{ -\frac{1}{2\sigma_W^2} \sum_{i=1}^{N} (y_i - \mu - \mu_{w_i})^2 \right\} \\ \cdot \quad & \frac{1}{\sigma_B^M} \exp\left\{ -\frac{1}{2\sigma_B^2} \sum_{j=1}^{M} \mu_j^2 \right\} \\ \cdot \quad & f(\sigma_W^2) f(\sigma_B^2) I_\mu(\mu_0, \mu_1) \, . \end{aligned}$$

It follows that the conditional posterior density for worker k's mean is propotional to

$$\begin{aligned} f(\mu_k|y,\mu,\sigma_W,\sigma_B) &\propto \exp\left\{-\frac{1}{2\sigma_W^2}\sum_{i:w_i=k}(y_i-\mu-\mu_k)^2\right\}\exp\left\{-\frac{1}{2\sigma_B^2}\mu_k^2\right\} \\ &= \exp\left\{-\frac{1}{2\sigma_W^2}\left(\sum_{i:w_i=k}(y_i-\mu)^2-2\sum_{i:w_i=k}(y_i-\mu)\mu_k+\mu_k^2 n_k\right)\right\}\exp\left\{-\frac{1}{2\sigma_B^2}\mu_k^2\right\}, \end{aligned}$$

where n_k is the number of observations in worker k. Thus

$$f(\mu_k|y,\mu,\sigma_W,\sigma_B) \propto \exp\left\{-\frac{n_k}{2\sigma_W^2}\left(\mu_k^2-2(\bar{y_k}-\mu)\mu_k\right)\right\}\exp\left\{-\frac{1}{2\sigma_B^2}\mu_k^2\right\},$$

where $\bar{y_k}$ is the average of observations for worker k.

Simple algebra reduces the above expression to

$$\mu_k|y,\mu,\sigma_W,\sigma_B \sim N\left(\frac{(\bar{y_k}-\mu)\sigma_B^2}{\sigma_W^2/n_k+\sigma_B^2}, \frac{\sigma_W^2\sigma_B^2/n_k}{\sigma_W^2/n_k+\sigma_B^2}\right). \tag{11}$$

The full conditional posterior density for μ is proportional to

$$\begin{aligned} f(\mu|y,\mu_1,\ldots,\mu_M,\sigma_W^2) &\propto \exp -\frac{1}{2\sigma_W^2}\left\{\sum_{i=1}^{N}(y_i-\mu-\mu_{w_i})^2\right\}I_\mu(\mu_0,\mu_1) \\ &= \exp -\frac{1}{2\sigma_W^2}\left\{\sum_i (y_i-\mu_{w_i})^2-2\mu\sum_i(y_i-\mu_{w_i})+N\mu^2\right\}I_\mu(\mu_0,\mu_1) \\ \Rightarrow \mu|y,\mu_1,\ldots,\mu_M,\sigma_W^2 &\sim N\left(\frac{\sum_i(y_i-\mu_{w_i})}{N},\frac{\sigma_W^2}{N}\right) \quad \text{truncated on } (\mu_0,\mu_1) \\ &\sim N\left(\bar{y}-\frac{\sum n_k\mu_k}{N},\frac{\sigma_W^2}{N}\right) \quad \text{truncated on } (\mu_0,\mu_1) \end{aligned} \tag{12}$$

The full conditional posterior density for σ_W is proportional to

$$\begin{aligned} f(\sigma_W|y,\mu,\mu_1,\ldots,\mu_M) &\propto \frac{1}{\sigma_W^N}\exp\left\{-\frac{1}{2\sigma_W^2}\sum_{i=1}^{N}(y_i-\mu-\mu_{w_i})^2\right\}\frac{1}{\sigma_W}\exp\left\{-\frac{1}{2\sigma^{*2}}\left(\log(\sigma_W)-\mu^*\right)^2\right\} \\ &= \frac{1}{\sigma_W^{N+1}}\exp\left\{-\frac{1}{2\sigma_W^2}\sum_{i=1}^{N}(y_i-\mu-\mu_{w_i})^2\right\}\exp\left\{-\frac{1}{2\sigma^{*2}}\left(\log(\sigma_W)-\mu^*\right)^2\right\} \end{aligned} \tag{13}$$

It follows that σ_W values can be sampled from the inverse cumulative density function (icdf) method presented in Appendix A, with

$$\begin{aligned} a &= N+1, \\ b &= \frac{1}{2}\sum_{i=1}^{N}(y_i-\mu-\mu_{w_i})^2, \\ \tilde{\mu} &= \mu^* \text{ and} \\ \tilde{\sigma}^2 &= \tilde{\sigma}^{*2}. \end{aligned}$$

Finally, the full conditional posterior density for σ_B^2 is proportional to

$$f(\sigma_B|y,\mu_1,\ldots,\mu_M) \propto \frac{1}{\sigma_B^M}\exp\left\{-\frac{1}{2\sigma_B^2}\sum_{j=1}^{M}\mu_j^2\right\}\frac{1}{\sigma_B}\exp\left\{-\frac{1}{2\sigma_B^{*2}}\left(\log(\sigma_B)-\mu_B^*\right)^2\right\}$$

Thus, σ_B can be sampled from its conditional posterior density through icdf method with

$$\begin{aligned} a &= M, \\ b &= \frac{1}{2}\sum_{j=1}^{M}\mu_j^2, \\ \tilde{\mu} &= \mu_B^* \text{ and} \\ \tilde{\sigma}^2 &= \sigma_B^{*2}\ . \end{aligned}$$

4.1 Alternative posterior when σ_W and σ_B prior distributions are uniform

An alternative to the above model is to use a uniform prior distribution on both σ_W and σ_B rather than the log-normal prior distributions specified in (9) and (10); the two *sigma* variables could possibly be defined on a specified range only, rather than on $\Re^+$.

When using uniform prior distribution on σ_B and σ_W, their respective full conditional posterior distributions are

$$\begin{aligned} f(\sigma_B|y,\mu_1,\ldots,\mu_M) &\propto \frac{1}{\sigma_B^M}\exp\left\{-\frac{1}{2\sigma_B^2}\sum_{j=1}^{M}\mu_j^2\right\} \\ \text{and } f(\sigma_W|y,\mu_1,\ldots,\mu_M) &\propto \frac{1}{\sigma_W^N}\exp\left\{-\frac{1}{2\sigma_W^2}\sum_{i=1}^{N}\left(y_i-\mu-\mu_{w_i}\right)^2\right\} \end{aligned} \tag{14}$$

from which we can easily sample using the algorithm described in Appendix A if M(or N) ≤ 1, or from an Inverted-Gamma distribution otherwise.

5 Banerjee model [SEG.riskband]

In the Banerjee model, the outcome follows either a normal distribution

$$Y_i \sim N(\mu,\sigma^2), i = 1,2,\ldots,N$$

or a lognormal distribution, given parameters (μ,σ).

Prior knowledge is expressed around the 95th percentile of the data distribution, that is, $Y_{0.95} = \mu + 1.645\sigma$; a series of cut-offs points $A_1, A_2, \ldots, A_{R-1}$ is given along with prior probabilities that (μ,σ) falls in either of the regions

$$\begin{aligned} \mathcal{R}_1 &= \{(\mu,\sigma) : \mu + z\sigma \leq A_1\} \\ \mathcal{R}_2 &= \{(\mu,\sigma) : A_1 < \mu + z\sigma \leq A_2\} \\ &\vdots \\ \mathcal{R}_{R-1} &= \{(\mu,\sigma) : A_{R-2} < \mu + z\sigma \leq A_{R-1}\} \\ \mathcal{R}_R &= \{(\mu,\sigma) : \mu + z\sigma > A_{R-1}\} \end{aligned}$$

with probabilities

$$P\{(\mu,\sigma) \in \mathcal{R}_j\} = p_j, \quad j = 1,2,\ldots,R \tag{15}$$

where z is the desired quantile of the normal distribution, in general the 95th, that is, $z = 1.645$.

We assume a piecewise-uniform joint prior distribution for (μ,σ) on its rectangle domain $(\mu_0,\mu_1)\times(\sigma_0,\sigma_1)$, where (μ_0,μ_1) and (σ_0,σ_1) are the lower and upper limits for μ and σ respectively, that is

$$f(\mu,\sigma) = \begin{cases} f_1 \text{ for } (\mu,\sigma) \in \mathcal{R}_1 \\ f_2 \text{ for } (\mu,\sigma) \in \mathcal{R}_2 \\ \vdots \\ f_R \text{ for } (\mu,\sigma) \in \mathcal{R}_R\ . \end{cases}$$

Given the clinical cut-off points $A_1, A_2, \ldots, A_{R-1}$, the domain of (μ, σ) is split into R regions. The figure below illustrates a typical case of the domain partitioning done that way.

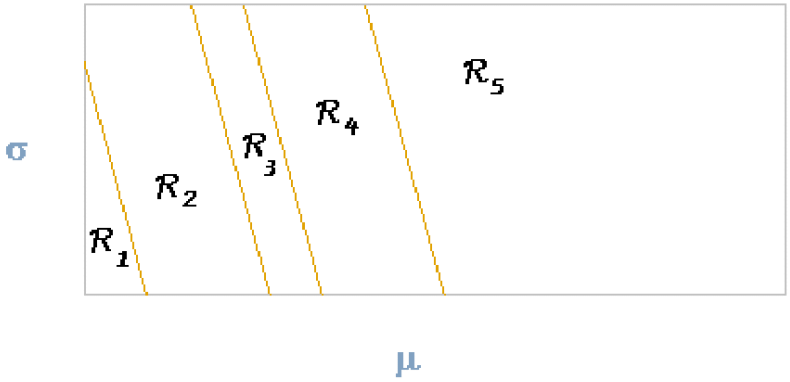


Figure 1: Regions defined by a series of clinical cut-off points.

If we let $S_j = \text{area}(\mathcal{R}_j), j = 1, 2, \ldots, R$, then the constant densities $f_1, f_2, \ldots, f_R$ for all points in each of the regions $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_R$ can be chosen such that

$$\begin{aligned} f_j &\geq 0, \quad j = 1, 2, \ldots, R, \\ \sum f_j &= 1 \\ \text{and } f_j \times S_j &\propto p_j, \quad j = 1, 2, \ldots, R. \end{aligned}$$

That is easily done by setting $f'_1 = 1$ and f'_j such that

$$\frac{f'_j S_j}{f'_1 S_1} = \frac{p_j}{p_1}, j \geq 2$$

and then $f_j = f'_j / \sum_k f'_k S_k, j = 1, 2, \ldots, R$.

The posterior distribution of (μ, σ) is thus

$$f(\mu, \sigma | y) \propto \frac{1}{\sigma^N} \exp\left\{-\frac{1}{2\sigma^2} \sum_i (y_i - \mu)^2\right\} f(\mu, \sigma) \, .$$

The full conditional posterior distribution for μ is proportional to

$$\begin{aligned} f(\mu | y, \sigma) &\propto \exp\left\{-\frac{1}{2\sigma^2}\left(N\mu^2 - 2\mu \sum y_i + \sum y_i^2\right)\right\} f(\mu | \sigma) \\ \Rightarrow \mu | y, \sigma &\sim N\left(\bar{y}, \frac{\sigma^2}{N}\right) f(\mu | \sigma) \, . \end{aligned}$$

Therefore, the full conditional posterior distribution for μ is a Normal density with sections defined by the μ-partition brought by $f(\mu|\sigma)$ weighed with the corresponding weights $f(\mu|\sigma)$.

In an attempt to clarify the result above, consider the typical figure below, suppose that $\sigma = \sigma^*$ and that we want to sample from $f(\mu | y, \sigma^*)$.

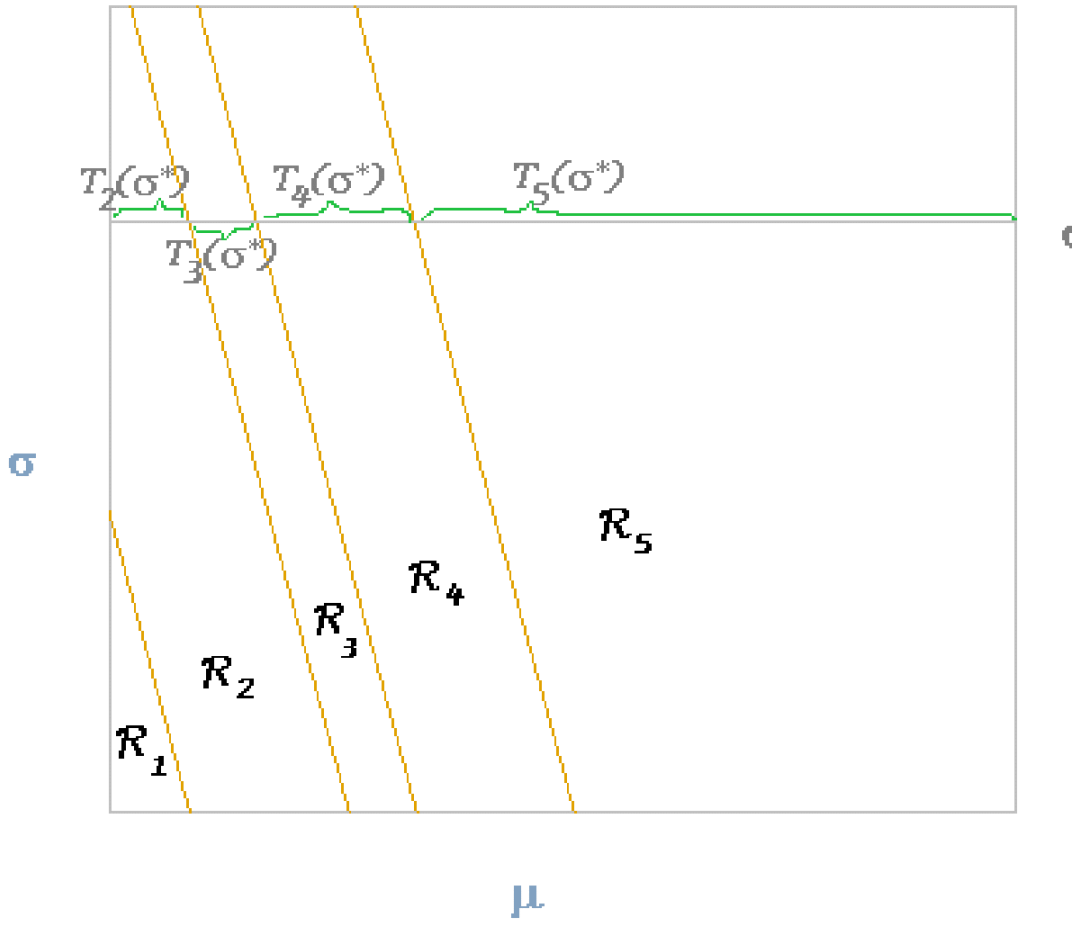


Figure 2: Partitioning μ-domain for a value $\sigma = \sigma^*$.

For the values $\mu \in (\mu_0, \mu_1)$, the couple (μ, σ^*) falls on either of the segments $T_2(\sigma^*)$, $T_3(\sigma^*)$, $T_4(\sigma^*)$ or $T_5(\sigma^*)$ and the density for the points in each segment $T_j(\sigma^*)$ is f_j, $j = 2, \ldots, 5$. Hence

$$f(\mu|\sigma^*) = f_j \text{ for } \mu \in T_j(\sigma^*),\ j = 2, \ldots, R \tag{16}$$

and the sampling of a value from μ conditional posterior distribution $f(\mu|y, \sigma^*)$ is trivial.

The full conditional posterior distribution for σ is given by

$$f(\sigma|y, \mu) \propto \frac{1}{\sigma^N} \exp\left\{-\frac{1}{2\sigma^2}\sum_i (y_i - \mu)^2\right\} f(\sigma|\mu)\ ,$$

that is, we can sample σ values from its conditional posterior distribution using the algorithm presented in Appendix A with

$$\begin{aligned} a &= N \text{ and} \\ b &= \frac{1}{2}\sum_i (y_i - \mu)^2, \end{aligned}$$

with sections defined through $f(\sigma|\mu)$, weighed by the corresponding f-weights. The partition emerging from $f(\sigma|\mu)$ can be derived the same way as μ-partition was derived from $f(\mu|\sigma)$ above.

5.1 Problems with the algorithm suggested by Banerjee

The model and the algorithm described in previous section follow the idea described in Banerjee's paper, but NOT the algortihm he suggested.

Erdogan Gunel raised the fact that the algorithm suggested by Banerjee does not provide a sample from the different regions $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_R$ with probabilities $p_1, p_2, \ldots, p_R$ as expected. Even though his demonstration — based on simulation — is incorrect (we reproduced his results by assuming an incorrect understanding of Banerjee's algorithm) and exacerbates the shift between observed proportions of pairs (μ, σ) in each region $\mathcal{R}_1$ to $\mathcal{R}_R$ and the corresponding proportions p_1 to p_R, he is right: the algorithm suggested by Banerjee does not provide a prior sample from each region with the expected probabilities p_1 to p_R.

Indeed, if the joint prior distribution for (μ, σ) is constant within each region R_j, then the marginal prior distribution for σ is **not** uniform — unless the region $\mathcal{R}_j$ is paralleliped-shaped, which is not the case for all regions (see regions $\mathcal{R}_1, \mathcal{R}_2$ and $\mathcal{R}_5$ in Figure 1); it is particularly easy to realize when considering the triangle-shaped region $\mathcal{R}_1$ from that Figure: consider $\sigma_1 > \sigma_0$ both falling in that triangle. The set of values $(\mu, \sigma_1) \in \mathcal{R}_1$ is clearly smaller than the set $(\mu, \sigma_0) \in \mathcal{R}_1$, and hence $f(\sigma_1) < f(\sigma_0)$, which shows that the marginal distribution for σ is NOT uniform; however, the algorithm suggested by Banerjee uses a uniform distribution as the marginal distribution for σ, which is clearly incorrect.

Writing a correct model in either WinBUGS or RJags for Banerjee's model is made difficult by the need to write a piecewise-uniform prior distribution for (μ, σ). Writing a Gibbs model (in R) based on repetitive sampling from the full conditional distributions $f(\mu|y, \sigma)$ and $f(\sigma|y, \mu)$ removes that difficulty as we have seen in the previous section.

The algorithm in the previous section also addresses the incorrect piecewise-density function suggested by Banerjee. Indeed, if $f(\mu, \sigma) = p_j$ for $(\mu, \sigma) \in \mathcal{R}_j$, as Banerjee suggests, then the prior probabilities for each region is $P\{(\mu, \sigma) \in \mathcal{R}_j\} \propto f_j \times S_j$, which can be far off the expected probabilites $p_1, p_2, \ldots, p_R$ when the areas $S_1, S_2, \ldots, S_R$ are dissimilar.

Finally, a word on the model suggested by Gunel: the author suggests a normal prior distribution for μ and an inverted-gamma distribution for σ, which can differ substantially from Banerjee's model. More importantly, the model does not include any prior knowledge about the probabilities of (μ, σ) in each region, even though the author seemed to consider the lack of control of the latter as the main caveat of Banerjee's model; hence it can hardly be presented as an alternative to Banerjee's model. Finally, the inverted-gamma prior distribution on σ decreases the model's value in practice.

5.2 Implementation in RJags

In an RJags version of Banerjee's algorithm, we need to sample (μ, σ) from their joint prior distribution $f(\mu, \sigma)$, presented in the above sections: that will be done through first sampling σ from its marginal prior distribution $f(\sigma)$ and then sampling μ from its conditional prior distribution $f(\mu|\sigma)$.

Suppose that the ranges for μ and σ and the cut-off values $A_1, A_2, \ldots, A_{R-1}$ are such that the full space for (μ, σ) corresponds to Figure 3.

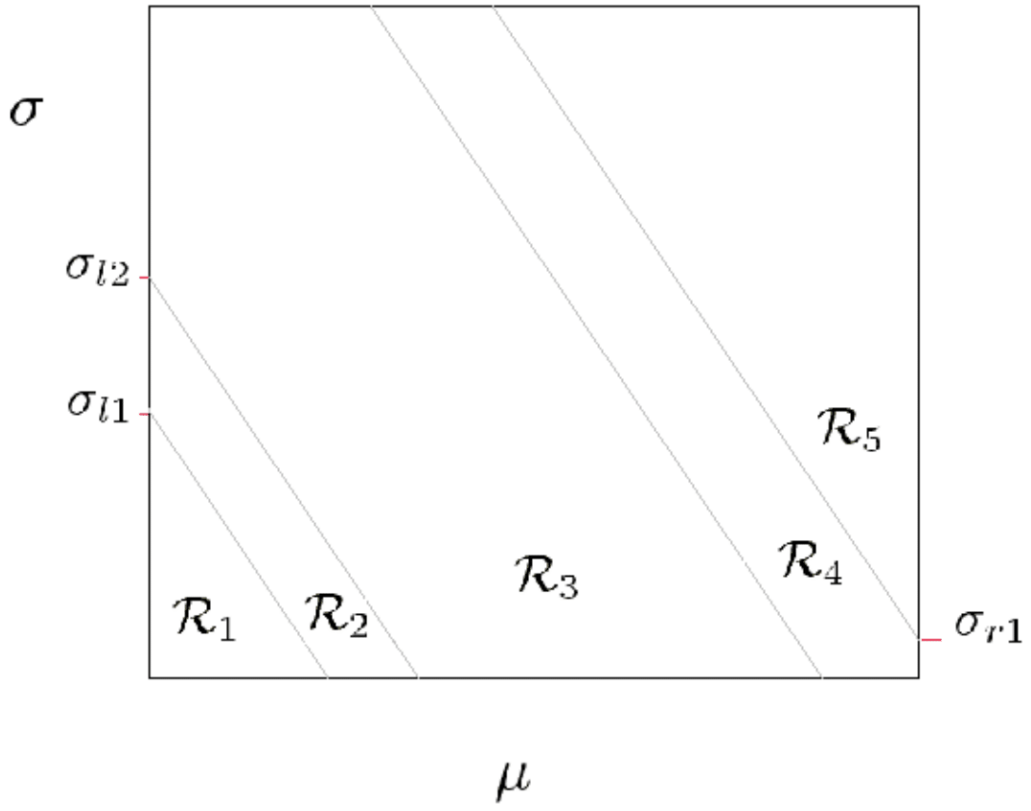


Figure 3: Regions and σ intercepts on left- and right-hand sides

It is easy to find the intercepts of the diagonal lines $\mu + z\sigma = A_i, i = 1, 2, \ldots, R$ with both the left-hand and the right-hand side borders of (μ, σ)'s domain — in the above figure, the values $(\sigma_{l1}, \sigma_{l2})$ and σ_{r1}, respectively. These values, once sorted, give the limits of J ($J = 4$ in this example) distinct intervals for σ, $\Sigma_1, \Sigma_2, \ldots, \Sigma_J$ such that $\cup\Sigma_j$ is equal to the whole domain for σ, as sketched in Figure 4.

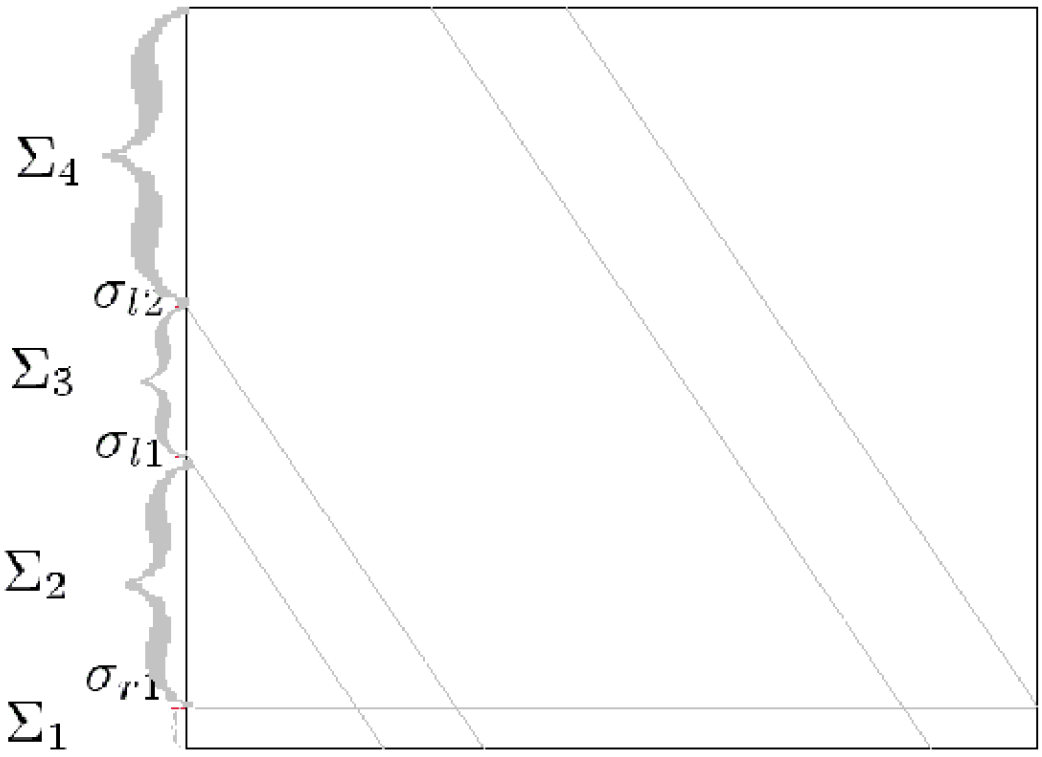


Figure 4: Segmenting σ's domain

The marginal density $f(\sigma)$ is easily calculated as

$$f(\sigma) = \sum_{j=1}^{R} l(T_j) f_j \tag{17}$$

where $l(T_j)$ is the length of the interval $T_j(\sigma)$, introduced in Figure 2. It is easy to see that $f(\sigma)$ is linear on each segment $\Sigma_1, \Sigma_2, \ldots, \Sigma_J$ and hence that the marginal prior distribution $f(\sigma)$ is piecewise-linear, as presented in Figure 5 below.

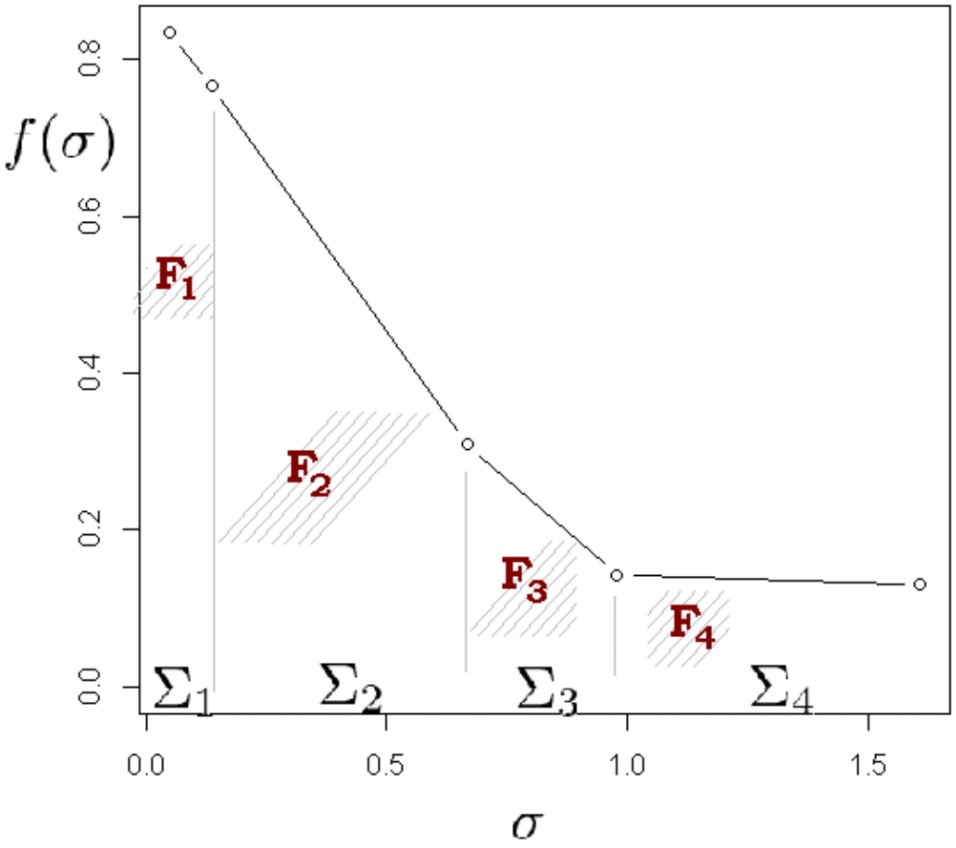


Figure 5: Piecewise-linear marginal prior density function $f(\sigma)$

Hence the marginal cumulative density function for σ, defined as $F(\sigma) = \int_{\sigma_0}^{\sigma} f(\sigma')d\sigma'$ is piecewise-quadratic. Let label Q_j the quadratic form of $F(\sigma)$ over the interval $\Sigma_j, j = 1, 2, \ldots, J$. Then one can sample σ through inverse cumulative density function by first sampling

$$U_\sigma \sim \text{Uniform}(0, 1)$$

and then solving

$$\sigma_j = Q_j^{-1}(U_\sigma)$$

for $j = 1, 2, \ldots, J$. The value σ_j is then accepted only if it is in the interval Σ_j: only one of $\{\sigma_1, \sigma_2, \ldots, \sigma_J\}$ is thus accepted, and the sample value for σ for this iteration is assigned the value of that unique solution.

Given a value sampled for σ, the conditional prior density for $f(\mu|\sigma)$ is trivially defined as

$$f(\mu|\sigma) = f_j \quad \text{if } \mu \in T_j(\sigma)$$

and sampling from it is also straightforward, as the conditional marginal cumulative density function for μ given σ is piecewise-linear. If we label L_j the linear equation describing $F(\mu|\sigma)$ over $T_j(\sigma), j = 1, 2, \ldots, R$, we can the sample a value μ by first sampling

$$U_\mu \sim \text{Uniform}(0, 1)$$

and then solving

$$\mu_j = L_j^{-1}(U_\mu)$$

for $j = 1, 2, \ldots, R$. The value μ_j is then accepted only if it is in the interval $T_j(\sigma)$: only one of $\{\mu_1, \mu_2, \ldots, \mu_R\}$ is thus accepted, and the sample value for μ for this iteration is assigned the value of that unique solution.

6 Modification of the algorithms in the presence of measurement error

The previous sections address the different models in the absence of measurement error. When measurement error exists, however, all need little adjustments. In the next two sections, we introduce two measurement error models. We first introduce the classical measurement error, assuming that the sd of the measurement error is the same for each observation — a model that appears to be of little interest in the context of Industrial Hygiene. In Section 6.2, we will thus introduce a second measurement error model where the scale of the measurement error is specified through a coefficient of variation ν, that is, where the sd of the measurement error is assumed to be some percentage of the (unobserved) true value.

In both measurement error models, the true values $T_i, i = 1, \ldots, N$ are latent variables. Depending on the model, T_i is assumed a Normal or a log-Normal distribution, that is, the likelihood is either

$$f(T_i|\mu,\sigma) = \begin{cases} f_1(T_i,\mu,\sigma) = \frac{1}{\sigma}\exp\left\{-\frac{1}{2\sigma^2}(T_i-\mu)^2\right\} & \text{when } T_i \sim N(\mu,\sigma^2) \\ f_2(T_i,\mu,\sigma) = \frac{1}{T_i\sigma}\exp\left\{-\frac{1}{2\sigma^2}(\log(T_i)-\mu)^2\right\} & \text{when } T_i \sim \text{logN}(\mu,\sigma^2) \end{cases} \tag{18}$$

6.1 Classical measurement error model

In the classical measurement error model, we assume that the observed values Y_i are normally distributed around the corresponding true values T_i with a standard deviation ξ independent of true value T_i and for which we assume a uniform prior distribution

$$\xi \sim \text{U}(\xi_0, \xi_1)\ .$$

Given the parameters (μ, σ) from the density of T_i, this measurement error model assumes that

$$Y_i|T_i \sim N(T_i, \xi^2),$$

that is,

$$f(Y_i|T_i,\xi) = g_1(Y_i,T_i,\xi) \propto \frac{1}{\xi}\exp\left\{-\frac{1}{2\xi^2}(Y_i-T_i)^2\right\}. \tag{19}$$

The full conditional posterior distribution for ξ is

$$f(\xi|Y,T) \propto \frac{1}{\xi^N}\exp\left\{-\frac{1}{\xi^2}\underbrace{\frac{1}{2}\sum_i (Y_i-T_i)^2}_{\beta}\right\} I_\xi(\xi_0,\xi_1)\ .$$

.

Hence we can easily sample from ξ's posterior distribution through a (truncated) Inverted-Gamma distribution if $N > 1$, or the icdf algorithm presented in Appendix A when $N \leq 1$.

6.1.1 Modification when the outcome is log-normally distributed

When the true values T_i are assumed to be log-normally distributed, their values will be positive. Consequently, assuming that the observed values Y_i are normally distributed around the latent true values seems unnatural, since it leads to non-zero probabilities of getting negative measured values; hence we propose the use of a normal distribution (around true values) *restricted to the positive domain* when the outcome is log-normally distributed.

In this context, the data distribution (19) being restricted to values $Y_i > 0$ must be divided by the standardizing constant $\kappa = \Pr(Y_i > 0) = \Phi(T_i/\xi)$ to integrate to 1; that is, we have

$$f(Y_i|T_i,\xi) = g_1^*(Y_i,T_i,\xi) \propto \frac{1}{\xi}\exp\left\{-\frac{1}{2\xi^2}(Y_i-T_i)^2\right\}\cdot\frac{1}{\Phi\left(\frac{T_i}{\xi}\right)} \tag{20}$$

and the full conditional posterior distribution for ξ must be modified accordingly, leading to

$$f(\xi|Y,T) \propto \frac{1}{\xi^N\prod_i\Phi\left(\frac{T_i}{\xi}\right)}\exp\left\{-\frac{1}{\xi^2}\underbrace{\frac{1}{2}\sum_i(Y_i-T_i)^2}_{\beta}\right\} I_\xi(\xi_0,\xi_1)\ . \tag{21}$$

Sampling values from the above distribution requires the use of the inverse cumulative density function method; the computation of the first two derivatives of $\log(f)$ necessary to do so is relegated to Section 8.

6.2 Measurement error specified through a coefficient of variation

In this measurement error model, we assume that the values Y_i are normally distributed around the corresponding true values T_i with a standard deviation equal to some percentage of the true value; that is, we assume a coefficient of variation ν known within some relatively close lower and upper limits ν_0 and ν_1, respectively, assume a uniform prior distribution

$$\nu \sim \mathrm{U}(\nu_0,\nu_1)$$

— e.g. $\nu \sim \mathrm{U}(15\%, 17\%)$ — and assume that the measurements Y_i are normally distributed around the true values T_i with standard deviation equal to $\nu \times T_i$, that is

$$Y_i|T_i \sim N(T_i, \mathrm{sd} = \nu T_i). \tag{22}$$

Before continuing on the elicitation of the full posterior distrbutions for parameter ν, we must emphasize on the latter making sense if and only if the values of $T_i, i = 1,2,\ldots,N$ are positive.

Warning: a restriction when the outcome is normally distributed. The data conditional (on the true values) distribution (22) makes sense if the outcome is log-normally distributed; however, if the data are normally distrbuted, there is a non-null probability that $T_i < 0$ and hence a non-null probability that $\mathrm{sd}(Y_i|T_i) = \nu T_i < 0$, which is obviously problematic! A work-around is to assume that the true values are **not** normally distributed, but rather normally distributed with values restricted to its positive domain, that is,

$$\begin{aligned} f(t_i|\mu,\sigma) &\propto f_{N(\mu,\sigma^2)}(t_i)I(t_i>0) \\ &= \frac{1}{\kappa}\cdot\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{t_i-\mu}{\sigma}\right)^2\right\}I(t_i>0) \end{aligned} \tag{23}$$

where κ is a standardizing constant. In order to have f integrate to 1, we must set

$$\kappa = \Pr(t_i > 0) = \Phi\left(\frac{\mu}{\sigma}\right)\ ;$$

that being a function of μ and σ, the posterior distributions developed earlier for these two parameters clearly need to be adjusted: that work is relegated to Section 7.

Since the range for true values T_i is restricted to positive values, it seems natural to confine the measured values Y_i to the positive domain as well, and hence the conditional distribution of data Y_i given true values T_i, including the appropriate standardizing constant $\kappa = \Pr(Y_i > 0) = \Phi(T_i/T_i\nu) = \Phi(1/\nu)$, is given by

$$f(Y_i|T_i,\nu) = g_2(Y_i,T_i,\nu) \propto \frac{1}{T_i\nu}\exp\left\{-\frac{1}{2T_i^2\nu^2}(Y_i-T_i)^2\right\}\cdot\frac{1}{\Phi(1/\nu)}. \tag{24}$$

The full conditional distribution for the coefficient of variation ν is proportional to

$$f(\nu|Y,T) \propto \frac{1}{\nu^N\Phi^N(1/\nu)}\exp\left\{-\frac{1}{\nu^2}\underbrace{\frac{1}{2}\sum_i\left(\frac{Y_i}{T_i}-1\right)^2}_{\beta}\right\} I_\nu(\nu_0,\nu_1)\ . \tag{25}$$

6.3 Conditional posterior distribution for T_i

In presence of measurement error, the full conditional posterior distribution for true value T_i is given by

$$f(T_i|Y_i,\mu,\sigma,\xi \text{ or } \nu) \propto f(Y_i|T_i,\xi \text{ or } \nu)f(T_i|\mu,\sigma^2) \tag{26}$$

which is equal to either $f_1g_1, f_1g_2, f_2g_1^*$ or f_2g_2 — described in (18), (19), (20) and (24) — depending on whether T_i is assumed to have a normal or a log-normal distribution and on the measurement error model.

The easiest case is when T_i is assumed to have a normal distribution and a classical measurement error model: in that scenario, the full posterior distribution for T_i is given by

$$\begin{aligned}
f(Y_i|T_i,\xi)f(T_i|\mu,\sigma^2) &\propto \exp\left\{-\frac{1}{2\xi^2}(Y_i-T_i)^2\right\}\exp\left\{-\frac{1}{2\sigma^2}(T_i-\mu)^2\right\}\\
&= \exp\left\{-\frac{1}{2\xi^2}(Y_i^2-2Y_iT_i+T_i^2)\right\}\exp\left\{-\frac{1}{2\sigma^2}(T_i^2-2T_i\mu+\mu^2)\right\}\\
&\propto \exp-\frac{1}{2}\left\{T_i^2\underbrace{\left(\frac{1}{\xi^2}+\frac{1}{\sigma^2}\right)}_{\tau^*} - 2T_i\underbrace{\left(\frac{Y_i}{\xi^2}+\frac{\mu}{\sigma^2}\right)}_{\mu_i^*}\right\}\\
&= \exp-\frac{\tau^*}{2}\left\{T_i^2-2T_i\frac{\mu_i^*}{\tau^*}\right\}\\
\Longrightarrow T_i|Y_i &\sim N\left(\frac{\mu_i^*}{\tau^*}, \text{precision}=\tau^*\right)\ .
\end{aligned} \tag{27}$$

6.3.1 When the measurement error is specified through a Coefficient of Variation and the outcome is log-normally distributed

The other three scenarios lead to conditional posterior distributions for T_i that are similar to each other but more complex than the above. For example, in the context of an outcome that is log-normally distributed and with a measurement error specified through a coefficient of variation (see Section 6.2), the full conditional posterior distribution for T_i is given by

$$\begin{aligned}
f(T_i|Y_i,\nu,\mu,\sigma) \propto f_2(T_i,\mu,\sigma,\nu,Y_i) &\propto \frac{1}{\nu T_i}\exp\left\{-\frac{1}{2\nu^2T_i^2}(Y_i-T_i)^2\right\}\cdot\frac{1}{T_i}\exp\left\{-\frac{1}{2\sigma^2}(\log(T_i)-\mu)^2\right\}\\
&\propto \frac{1}{T_i^2}\exp\left\{-\frac{1}{2\nu^2T_i^2}(Y_i-T_i)^2\right\}\exp\left\{-\frac{1}{2\sigma^2}(\log(T_i)-\mu)^2\right\}\ .
\end{aligned}$$

As one can see from the above, the full conditional distribution $f(T_i|Y_i,\nu,\mu,\sigma)$ is quite complex and there does not seem to be an easy way to sample from it; however, given values Y_i,μ,σ and ν, one can numerically evaluate the corresponding cumulative density and thus use the inverse cumulative density function algorithm to sample a random value from it. Sampling from T_i's posterior distribution in the other two scenarios is performed along the same lines.

Initial values for Y_i's are generated with a Normal or log-Normal distribution with parameters μ and σ^2 and limited to their corresponding censoring range.

At each following iteration, the values $T_i, i = 1,\ldots,N$ are generated from an icdf algorithm as mentioned above and then the censored values in $\{Y_i\}_{i=1}^N$ — if any — are generated with distributions (19) or (24) and limited to their corresponding censoring range.

6.3.2 When the measurement error is specified through a Coefficient of Variation and the outcome is normally distributed

When the outcome is normally distributed and the measurement error is specified through a Coefficient of Variation, the full conditional posterior distribution for T_i is given by

$$\begin{aligned} f(T_i|Y_i,\nu,\mu,\sigma) &\propto \frac{1}{T_i\nu\Phi(1/\nu)}\exp\left\{-\frac{1}{2T_i^2\nu^2}(Y_i-T_i)^2\right\}\exp\left\{-\frac{1}{2}\left(\frac{T_i-\mu}{\sigma}\right)^2\right\} \\ &\propto \frac{1}{T_i}\exp\left\{-\frac{1}{2T_i^2\nu^2}(Y_i-T_i)^2\right\}\exp\left\{-\frac{1}{2}\left(\frac{T_i-\mu}{\sigma}\right)^2\right\} \end{aligned}$$

for $T_i > 0$, from (23) and (24).

6.3.3 When the measurement error is classical and the outcome is log-normally distributed

Under the classical measurement error model and when the outcome is log-normally distributed, the full conditional posterior distribution for T_i is given by

$$\begin{aligned} f(T_i|Y_i,\xi,\mu,\sigma) &\propto \frac{1}{\xi}\exp\left\{-\frac{1}{2\xi^2}(Y_i-T_i)^2\right\}\frac{1}{\Phi\left(\frac{T_i}{\xi}\right)}\cdot\frac{1}{T_i\sigma}\exp\left\{-\frac{1}{2\sigma^2}(\log(T_i)-\mu)^2\right\} \\ &\propto \frac{1}{T_i\Phi\left(\frac{T_i}{\xi}\right)}\exp\left\{-\frac{1}{2\xi^2}(Y_i-T_i)^2\right\}\exp\left\{-\frac{1}{2\sigma^2}(\log(T_i)-\mu)^2\right\} \end{aligned} \tag{28}$$

from (18) and (20).

7 Revisiting posterior distributions in the presence of Measurement Error specified through a Coefficient of Variation when the outcome is normally distributed

As we have discussed in Section 6.2, the full conditional posterior distributions for μ,σ (and μ_k,σ_W in McNally's model) need to be revisited in presence of measurement error specified through a coefficient of variation when the outcome is normally distributed. The sections below will address these issues one model at a time, taking advantage of results developed in Appendix C.

Each of the full conditional posterior distributions developed below involve a term $\Phi\left(\frac{\mu}{\sigma}\right)$ in the denominator, making the sampling of random values from them impossible in a direct manner: we will rather use an inverse cumulative density function. For that algorithm to be efficient, we need to find the domain on which the conditional posterior density function is not negligeable, which will be done by first finding its mode and then remote left and right values (such that the density at these two endpoints is very small when compared to density at the mode). Hence we need to compute the first two derivatives of the log of each full conditional posterior distribution.

7.1 Uninformative model

The full conditional posterior distribution for μ (2) becomes

$$\begin{aligned} f &\propto \frac{e^{-\frac{N}{2\sigma^2}(\mu-\bar{y})^2}}{\Phi^N\left(\frac{\mu}{\sigma}\right)} \qquad (29)\\ \text{therefore } \log(f) &= C - \frac{N}{2\sigma^2}(\mu-\bar{y})^2 - N\log\Phi\left(\frac{\mu}{\sigma}\right) \end{aligned}$$

where the constant C can be ignored (we will omit it in the $\log(f)$ expressions throughout the remainder of this section). The two terms in $\log f$ above correspond to the functions h_4 and h_3, respectively, described in (C.5), and we thus easily find its first two derivatives

$$\begin{aligned} \frac{\partial}{\partial\mu}\log f &= -\frac{N}{\sigma^2}(\mu-\bar{y}) - \frac{N\phi}{\Phi\sigma} \qquad (30)\\ \frac{\partial^2}{\partial\mu^2}\log f &= -\frac{N}{\sigma^2} + \frac{N\phi}{\sigma^2\Phi^2}\left(\frac{\mu\Phi}{\sigma}+\phi\right) \qquad (31)\\ \text{where } \phi &= \phi\left(\frac{\mu}{\sigma}\right)\\ \text{and } \Phi &= \Phi\left(\frac{\mu}{\sigma}\right)\ . \end{aligned}$$

If we let

$$\begin{aligned} z &= \frac{\mu}{\sigma}\\ \text{and } \varphi &= \frac{\phi}{\Phi} \end{aligned}$$

then results (30) and (31) simplify to

$$\begin{aligned} \frac{\partial}{\partial\mu}\log f &= -\frac{N}{\sigma^2}(\mu-\bar{y}) - \frac{N\varphi}{\sigma} \qquad (32)\\ \frac{\partial^2}{\partial\mu^2}\log f &= -\frac{N}{\sigma^2} + \frac{N}{\sigma^2}\left(\frac{z\varphi}{\sigma^2}+\varphi^2\right)\\ &= \frac{N}{\sigma^2}\left(z\varphi+\varphi^2-1\right)\ . \qquad (33) \end{aligned}$$

The full conditional posterior distribution for σ (3) becomes

$$\begin{aligned} f &\propto \frac{1}{\sigma^N}\exp\left\{-\frac{b}{\sigma^2}\right\}\frac{1}{\Phi^N\left(\frac{\mu}{\sigma}\right)},\\ \text{therefore } \log f &= -N\log\sigma - \frac{b}{\sigma^2} - N\log\Phi\left(\frac{\mu}{\sigma}\right) \end{aligned}$$

whose terms correspond to functions h_1, h_2 and h_3 in (C.5), respectively. We thus easily find

$$\begin{aligned} \frac{\partial}{\partial\sigma}\log f &= -\frac{N}{\sigma} + \frac{2b}{\sigma^3} + \frac{N\mu\phi}{\Phi\sigma^2}\\ \text{and } \frac{\partial^2}{\partial\sigma^2}\log f &= \frac{N}{\sigma^2} - \frac{6b}{\sigma^4} + \frac{N\mu\phi}{\Phi^2\sigma^4}\left\{\frac{\mu^2}{\sigma}\Phi + \mu\phi - 2\sigma\Phi\right\}\ ; \end{aligned}$$

in terms of z and φ, the above two reduce to

$$\begin{aligned} \frac{\partial}{\partial\sigma}\log f &= -\frac{N}{\sigma} + \frac{2b}{\sigma^3} + \frac{Nz\varphi}{\sigma}\\ \text{and } \frac{\partial^2}{\partial\sigma^2}\log f &= \frac{N}{\sigma^2} - \frac{6b}{\sigma^4} + \frac{N\mu\phi}{\Phi^2\sigma^4}\left\{\frac{\mu^2}{\sigma}\Phi + \mu\phi - 2\sigma\Phi\right\}\\ &= \frac{N}{\sigma^2} - \frac{6b}{\sigma^4} + \frac{Nz}{\sigma^2}\left((z^2-2)\varphi + z\varphi^2\right). \end{aligned}$$

7.2 Kromhout model

The full conditional posterior distribution for μ (6) rewrites as (29) and therefore its first two derivatives are given by (32) and (33).

The full conditional posterior density for σ (7) rewrites

$$\begin{aligned} f &\propto \frac{1}{\sigma^{N+1}} \exp\left\{-\frac{b}{\sigma^2}\right\} \exp\left\{-\frac{(\log(\sigma)-\mu^*)^2}{2\sigma^{*2}}\right\} \frac{1}{\Phi^N\left(\frac{\mu}{\sigma}\right)} \\ \text{and } \log f &= -(N+1)\log\sigma - \frac{b}{\sigma^2} - \frac{(\log(\sigma)-\mu^*)^2}{2\sigma^{*2}} - N\log\Phi\left(\frac{\mu}{\sigma}\right) \end{aligned}$$

which terms correspond to functions h_1, h_2, h_5 and h_3 in (C.5), respectively.

Hence we easily get its first two derivatives, given by

$$\begin{aligned} \frac{\partial}{\partial\sigma}\log f &= -\frac{N+1}{\sigma} + \frac{2b}{\sigma^3} - \frac{1}{\sigma\sigma^{*2}}(\log(\sigma)-\mu^*) + \frac{N\mu\phi}{\Phi\sigma^2} \\ \text{and } \frac{\partial^2}{\partial\sigma^2}\log f &= \frac{N+1}{\sigma^2} - \frac{6b}{\sigma^4} + \frac{\log(\sigma)-\mu^*-1}{\sigma^{*2}\sigma^2} + \frac{N\mu\phi}{\Phi^2\sigma^4}\left\{\frac{\mu^2}{\sigma}\Phi + \mu\phi - 2\sigma\Phi\right\}; \end{aligned}$$

in terms of z and φ, the above two reduce to

$$\begin{aligned} \frac{\partial}{\partial\sigma}\log f &= -\frac{N+1}{\sigma} + \frac{2b}{\sigma^3} - \frac{1}{\sigma\sigma^{*2}}(\log(\sigma)-\mu^*) + \frac{Nz\varphi}{\sigma} \\ \text{and } \frac{\partial^2}{\partial\sigma^2}\log f &= \frac{N+1}{\sigma^2} - \frac{6b}{\sigma^4} + \frac{\log(\sigma)-\mu^*-1}{\sigma^{*2}\sigma^2} + \frac{Nz}{\sigma^2}\left((z^2-2)\varphi + z\varphi^2\right). \end{aligned}$$

7.3 McNally model

The full conditional posterior density for μ (12) becomes

$$\begin{aligned} f &\propto \frac{\exp\left\{-\frac{N}{2\sigma_W^2}(\mu-\theta)^2\right\}}{\prod_k \Phi^{n_k}\left(\frac{\mu+\mu_k}{\sigma_W}\right)} \quad \text{where } \theta = \bar{y} - \frac{\sum n_k\mu_k}{N} \\ \text{therefore } \log f &= -\frac{N}{2\sigma_W^2}(\mu-\theta)^2 - \sum_k n_k \log\Phi\left(\frac{\mu+\mu_k}{\sigma_W}\right) \end{aligned}$$

which terms correspond to h_4 and h_3 in (C.5), respectively. Its first two derivatives are easily obtained as

$$\begin{aligned} \frac{\partial}{\partial\mu}\log f &= -\frac{N}{\sigma_W^2}(\mu-\theta) - \frac{1}{\sigma_W}\sum_k \frac{n_k\phi_k}{\Phi_k} \\ \text{and } \frac{\partial^2}{\partial\mu^2}\log f &= -\frac{N}{\sigma_W^2} + \sum_k \frac{n_k\phi_k}{\sigma_W^2\Phi_k^2}\left(\frac{(\mu+\mu_k)\Phi_k}{\sigma_W} + \phi_k\right) \\ \text{where } \phi_k &= \phi\left(\frac{\mu+\mu_k}{\sigma_W}\right) \\ \text{and } \Phi_k &= \Phi\left(\frac{\mu+\mu_k}{\sigma_W}\right). \end{aligned}$$

If we let

$$\begin{aligned} z_k &= \frac{\mu+\mu_k}{\sigma_W} \\ \text{and } \varphi_k &= \frac{\phi_k}{\Phi_k} \end{aligned}$$

then the above two derivatives reduce to

$$\begin{aligned}\frac{\partial}{\partial\mu}\log f &= -\frac{N}{\sigma_W^2}(\mu-\theta)-\frac{1}{\sigma_W}\sum_k n_k\varphi_k \\ \text{and } \frac{\partial^2}{\partial\mu^2}\log f &= -\frac{N}{\sigma_W^2}+\frac{1}{\sigma_W^2}\sum_k n_k(z_k\varphi_k+\varphi_k^2)\ .\end{aligned}$$

The full conditional posterior density for μ_k (11) becomes

$$\begin{aligned}f &\propto \frac{\exp\left\{-\frac{1}{2\sigma_k^{*2}}(\mu_k-\mu_k^*)^2\right\}}{\Phi^{n_k}\left(\frac{\mu+\mu_k}{\sigma_W}\right)} \\ \text{therefore } \log f &= -\frac{1}{2\sigma_k^{*2}}(\mu_k-\mu_k^*)^2-n_k\log\Phi\left(\frac{\mu+\mu_k}{\sigma_W}\right)\end{aligned}$$

where μ_k^* and σ_k^{*2} are the mean and variance of the distribution in (11). The components of $\log f$ above correspond to h_4 and h_3 in (C.5), respectively. Its first two derivatives are easily obtained as

$$\begin{aligned}\frac{\partial}{\partial\mu_k}\log f &= -\frac{1}{\sigma_k^{*2}}(\mu_k-\mu_k^*)-\frac{n_k\phi_k}{\Phi_k\sigma_W} \\ \text{and } \frac{\partial^2}{\partial\mu_k^2}\log f &= -\frac{1}{\sigma_k^{*2}}+\frac{n_k\phi_k}{\sigma_W^2\Phi_k^2}\left(\frac{(\mu+\mu_k)\Phi_k}{\sigma_W}+\phi_k\right)\ ;\end{aligned}$$

in terms of z_k and φ_k, the above two reduce to

$$\begin{aligned}\frac{\partial}{\partial\mu_k}\log f &= -\frac{1}{\sigma_k^{*2}}(\mu_k-\mu_k^*)-\frac{n_k\varphi_k}{\sigma_W} \\ \text{and } \frac{\partial^2}{\partial\mu_k^2}\log f &= -\frac{1}{\sigma_k^{*2}}+\frac{n_k}{\sigma_W^2}(z_k\varphi_k+\varphi_k^2)\ .\end{aligned}$$

When the prior distribution for σ_W is log-normal, its posterior distribution (13) becomes

$$\begin{aligned}f &\propto \frac{1}{\sigma_W^{N+1}}\exp\left\{-\frac{b}{\sigma_W^2}\right\}\exp\left\{-\frac{1}{2\sigma^{*2}}\left(\log(\sigma_W)-\mu^*\right)^2\right\}\frac{1}{\prod_k\Phi^{n_k}\left(\frac{\mu+\mu_k}{\sigma_W}\right)} \\ \text{therefore } \log f &= -(N+1)\log\sigma_W-\frac{b}{\sigma_W^2}-\frac{1}{2\sigma^{*2}}\left(\log(\sigma_W)-\mu^*\right)^2-\sum_k n_k\log\Phi\left(\frac{\mu+\mu_k}{\sigma_W}\right)\end{aligned}$$

whose terms correspond to h_1, h_2, h_5 and h_3 in (C.5), respectively. Its first two derivatives are easily obtained as

$$\begin{aligned}\frac{\partial}{\partial\sigma_W}\log f &= -\frac{N+1}{\sigma_W}+\frac{2b}{\sigma_W^3}-\frac{1}{\sigma_W\sigma^{*2}}(\log(\sigma_W)-\mu^*)+\frac{1}{\sigma_W^2}\sum_k\frac{n_k(\mu+\mu_k)\phi_k}{\Phi_k} \\ \text{and } \frac{\partial^2}{\partial\sigma_W^2}\log f &= \frac{N+1}{\sigma_W^2}-\frac{6b}{\sigma_W^4}+\frac{\log(\sigma_W)-\mu^*-1}{\sigma^{*2}\sigma_W^2} \\ &\quad +\sum_k\frac{n_k(\mu+\mu_k)\phi_k}{\Phi_k^2\sigma_W^4}\left\{\frac{(\mu+\mu_k)^2}{\sigma_W}\Phi_k+(\mu+\mu_k)\phi_k-2\sigma_W\Phi_k\right\}\ ;\end{aligned}$$

in terms of z_k and φ_K, the above two reduce to

$$\begin{aligned}\frac{\partial}{\partial\sigma_W}\log f &= -\frac{N+1}{\sigma_W}+\frac{2b}{\sigma_W^3}-\frac{1}{\sigma_W\sigma^{*2}}(\log(\sigma_W)-\mu^*)+\frac{1}{\sigma_W}\sum_k n_k z_k\varphi_k \\ \text{and } \frac{\partial^2}{\partial\sigma_W^2}\log f &= \frac{N+1}{\sigma_W^2}-\frac{6b}{\sigma_W^4}+\frac{\log(\sigma_W)-\mu^*-1}{\sigma^{*2}\sigma_W^2}+\frac{1}{\sigma_W^2}\sum_k n_k z_k((z_k^2-2)\varphi_k+z_k\varphi_k^2)\ .\end{aligned}$$

When the prior distribution for σ_W is uniform, its posterior distribution (14) rewrites

$$\begin{aligned} f &\propto \frac{1}{\sigma_W^N} \exp\left\{-\frac{b}{\sigma_W^2}\right\} \frac{1}{\prod_k \Phi^{n_k}\left(\frac{\mu+\mu_k}{\sigma_W}\right)} \\ \text{therefore } \log f &= -N\log\sigma_W - \frac{b}{\sigma_W^2} - \sum_k n_k \log\Phi\left(\frac{\mu+\mu_k}{\sigma_W}\right) \end{aligned}$$

which terms correspond to h_1, h_2 and h_3 in (C.5), respectively. Its first two derivatives are easily obtained as

$$\begin{aligned} \frac{\partial}{\partial\sigma_W}\log f &= -\frac{N}{\sigma_W} + \frac{2b}{\sigma_W^3} + \frac{1}{\sigma_W^2}\sum_k \frac{n_k(\mu+\mu_k)\phi_k}{\Phi_k} \\ \text{and } \frac{\partial^2}{\partial\sigma_W^2}\log f &= \frac{N}{\sigma_W^2} - \frac{6b}{\sigma_W^4} + \sum_k \frac{n_k(\mu+\mu_k)\phi_k}{\Phi_k^2\sigma_W^4}\left\{\frac{(\mu+\mu_k)^2}{\sigma_W}\Phi_k + (\mu+\mu_k)\phi_k - 2\sigma_W\Phi_k\right\} ; \end{aligned}$$

in terms of z_k and φ_K, the above two reduce to

$$\begin{aligned} \frac{\partial}{\partial\sigma_W}\log f &= -\frac{N}{\sigma_W} + \frac{2b}{\sigma_W^3} + \frac{1}{\sigma_W}\sum_k n_k z_k\varphi_k \\ \text{and } \frac{\partial^2}{\partial\sigma_W^2}\log f &= \frac{N}{\sigma_W^2} - \frac{6b}{\sigma_W^4} + \frac{1}{\sigma_W^2}\sum_k n_k z_k\left\{(z_k^2-2)\varphi_k + z_k\varphi_k^2\right\} . \end{aligned}$$

7.4 Banerjee model

Since the posterior distributions for μ and σ under the Banerjee model are piecewise, the adjustments needed in the presence of measurement error specified through a coefficient of variation when the outcome is normally distributed are the same as for uninformative model (see Section 7.1).

7.5 Posterior distribution for ν

Whether the outcome is normally or log-normally distributed, the posterior distribution for the coefficient of variation ν (25) becomes

$$\begin{aligned} f &\propto \frac{1}{\nu^N\Phi^N\left(\frac{1}{\nu}\right)} \exp\left\{-\frac{\beta}{\nu^2}\right\} \\ \text{therefore } \log f &= -N\log\nu - \frac{\beta}{\nu^2} - N\log\Phi\left(\frac{1}{\nu}\right) \end{aligned}$$

whose terms correspond to functions h_1, h_2 and h_3 in (C.5), respectively. Its first two derivatives can thus be easily written as

$$\begin{aligned} \frac{\partial}{\partial\nu}\log f &= -\frac{N}{\nu} + \frac{2\beta}{\nu^3} + \frac{N\phi}{\Phi\nu^2} \\ \text{and } \frac{\partial^2}{\partial\nu^2}\log f &= \frac{N}{\nu^2} - \frac{6\beta}{\nu^4} + \frac{N}{\Phi^2\nu^4}\left\{\frac{\phi}{\nu^3}\Phi\nu^2 - \phi\left[-\frac{\phi}{\nu^2}\nu^2 + 2\Phi\nu\right]\right\} \\ &= \frac{N}{\nu^2} - \frac{6\beta}{\nu^4} + \frac{N\phi}{\Phi^2\nu^4}\left\{\frac{\Phi}{\nu} + \phi - 2\Phi\nu\right\} \\ \text{where } \phi &= \phi\left(\frac{1}{\nu}\right) \\ \text{and } \Phi &= \Phi\left(\frac{1}{\nu}\right) ; \end{aligned}$$

if we let $\varphi = \phi/\Phi$, then the above two results reduce to

$$\begin{aligned} \frac{\partial}{\partial\nu}\log f &= -\frac{N}{\nu} + \frac{2\beta}{\nu^3} + \frac{N\varphi}{\nu^2} \\ \text{and } \frac{\partial^2}{\partial\nu^2}\log f &= \frac{N}{\nu^2} - \frac{6\beta}{\nu^4} + \frac{N}{\nu^4}\left\{\varphi\left(\frac{1}{\nu} - 2\nu\right) + \varphi^2\right\} . \end{aligned}$$

8 Posterior distribution for classical measurement error parameter (ξ) when the outcome is log-normally distributed

Sampling values for ξ from its posterior distribution when the outcome is log-normally distributed (see equation 21) also requires the computation of the first two derivatives of $\log(f)$.

Since

$$\log(f) = -N\log(\xi) - \sum_i \log\Phi\left(\frac{T_i}{\xi}\right) - \frac{\beta}{\xi^2}$$

where the second term corresponds to h_3 in (C.5) — replacing $\mu+\theta$ by T_i and σ by ξ — we easily get

$$\frac{\partial}{\partial\xi}\log(f) = -\frac{N}{\xi} + \sum_i \frac{1}{\Phi\left(\frac{T_i}{\xi}\right)}\cdot\phi\left(\frac{T_i}{\xi}\right)\frac{T_i}{\xi^2} + \frac{2\beta}{\xi^3}$$

from derivation chain rule and

$$\begin{aligned} \frac{\partial^2}{\partial\xi^2}\log(f) &= \frac{N}{\xi^2} + \sum_i \frac{T_i\phi_i}{\Phi_i^2\xi^4}\left\{\frac{\Phi_i T_i^2}{\xi} + \phi_i - 2\xi\Phi_i\right\} - \frac{6\beta}{\xi^4} \\ &\quad \text{where } \phi_i = \phi\left(\frac{T_i}{\xi}\right) \text{ and } \Phi_i = \Phi\left(\frac{T_i}{\xi}\right) \\ &= \frac{N}{\xi^2} + \sum_i \frac{T_i\varphi_i}{\xi^5}\left\{T_i^2 + \xi\varphi_i T_i - 2\xi^2\right\} - \frac{6\beta}{\xi^4} \\ &\quad \text{where } \varphi_i = \frac{\phi_i}{\Phi_i}. \end{aligned}$$

A Generating values for σ from its inverse cumulative density function

If U is a random variable with a Uniform(0,1) density, then the variable $X = F^{-1}(U)$ has the cumulative density function F.

This method will be used in the context of WebExpo for σ when its conditional posterior distribution is given by either

$$f(\sigma|y, \text{other parameters}) \propto \frac{1}{\sigma^{a+1}} \exp\left\{-b/\sigma^2\right\} \exp\left\{-\frac{(\log(\sigma) - \tilde{\mu})^2}{2\tilde{\sigma}^2}\right\} ,$$

as is the case in the Kromhout, Two-Level Kromhout and McNally models, or

$$f(\sigma|y, \text{other parameters}) \propto \frac{1}{\sigma^{a}} \exp -b/\sigma^2 ,$$

as is the case in the Banerjee model and the uninformative model when the prior on σ is uniform.

In either case, the cumulative density function $F(\sigma) = \int_{-\infty}^{\sigma} f(\sigma')d\sigma'$ does not have an analytic solution but can be estimated numerically in R with the `integrate()` function for any value σ.

Hence, one can sample a value U from a uniform $U(0, 1)$ distribution and use a Newton-Raphson algorithm to find the value for σ such that

$$\frac{F(\sigma) - F(\sigma_0)}{F(\sigma_1) - F(\sigma_0)} = U$$

where (σ_0, σ_1) are the boundaries of the σ-domain; the resulting value σ is thus sampled from its corresponding f posterior density.

B Distribution of σ when precision τ is Gamma distributed

When the precision $\tau = 1/\sigma^2$ is Gamma-distributed, that is, when

$$\begin{aligned} \tau &\sim \text{Gamma}(\alpha, \beta) \\ \text{or } f(\tau) &\propto \tau^{\alpha-1} e^{-\beta\tau} \end{aligned}$$

then the distribution of σ is given by

$$f(\sigma) \propto \frac{1}{\sigma^{2(\alpha-1)}} \exp\left\{-\frac{\beta}{\sigma^2}\right\} \cdot \sigma^{-3}$$

since

$$\begin{aligned} \tau &= \sigma^{-2} \\ \text{and } \frac{d\tau}{d\sigma} &= -2\sigma^{-3}. \end{aligned}$$

Hence

$$\begin{aligned} f(\sigma) &\propto \frac{1}{\sigma^{2\alpha-2+3}} \exp\left\{-\frac{\beta}{\sigma^2}\right\} \\ &= \frac{1}{\sigma^{2\alpha+1}} \exp\left\{-\frac{\beta}{\sigma^2}\right\} \end{aligned} \tag{B.1}$$

Conversely, if

$$\begin{aligned} f(\sigma) &\propto \frac{1}{\sigma^a} \exp\left\{-\frac{\beta}{\sigma^2}\right\} \\ \text{then } \tau &\sim \text{Gamma}\left(\frac{a-1}{2}, \beta\right). \end{aligned} \tag{B.2}$$

C Derivatives

C.1 First derivative of ϕ and Φ

Let $\phi(\cdot)$ and $\Phi(\cdot)$ be the normal density and cumulative density functions, respectively, that is

$$\begin{aligned} \phi(z) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \\ \text{and } \Phi(z) &= \int_{-\infty}^{z} \phi(x)dx \, . \end{aligned}$$

If $u = u(x)$, then the first derivatives of $\phi = \phi(u(x))$ and $\Phi = \Phi(u(x))$ are

$$\begin{aligned} \phi' &= \frac{d}{dx}\phi(u(x)) \\ &= \frac{1}{\sqrt{2\pi}} \frac{d}{dx} e^{-\frac{u^2(x)}{2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2(x)}{2}} \cdot -\frac{2u(x)}{2} \cdot u' \\ &= -\phi u u' \end{aligned} \tag{C.1}$$

and

$$\begin{aligned} \Phi' &= \frac{d}{dx}\Phi(u(x)) \\ &= \phi(u(x)) \cdot \frac{d}{dx} u(x) \\ &= \phi u' \, . \end{aligned} \tag{C.2}$$

C.2 Derivatives of $\log(\Phi(g(z)))$

The first two derivatives of $\log(\Phi(g(z)))$ are given by

$$\begin{aligned} \frac{\partial}{\partial z} \log(\Phi(g(z))) &= \frac{1}{\Phi(g(z))} \cdot \phi(g(z))g'(z) \\ &\quad \text{from derivation chain rule} \\ &= \varphi(g(z))g'(z) \text{ where } \varphi(z) = \frac{\phi(z)}{\Phi(z)} \end{aligned} \tag{C.3}$$

and

$$\begin{aligned} \frac{\partial^2}{\partial z^2} \log(\Phi(g(z))) &= \frac{\partial}{\partial z}\left[\frac{\phi \cdot g'}{\Phi}\right] \\ &= \frac{[\phi g']'\Phi - \phi g' \Phi'}{\Phi^2} \\ &= \frac{(\phi' g' + \phi g'')\Phi - \phi g' \phi g'}{\Phi^2} \text{ from (C.2)} \\ &= \frac{\phi' g' + \phi g''}{\Phi} - \left[\frac{\phi g'}{\Phi}\right]^2 \\ &= \frac{-\phi g g'^2 + \phi g''}{\Phi} - \left[\frac{\phi g'}{\Phi}\right]^2 \text{ from (C.1)} \\ &= \frac{\phi}{\Phi}\left[-g g'^2 + g''\right] - \left[\frac{\phi g'}{\Phi}\right]^2 \\ &= \varphi\left[-g g'^2 + g''\right] - \left[\frac{\phi g'}{\Phi}\right]^2 . \end{aligned} \tag{C.4}$$

C.3 Derivatives of functions involved in full conditional posterior distributions

The posterior distributions elicited in Section 7 are the product of density functions (from likelihood and prior distributions) which can be regrouped in 5 types of functions, which we label $h1, \ldots, h_5$, defined as

$$h_i = \begin{cases} \frac{1}{\sigma^a} & \text{when } i = 1 \\ \exp -\frac{\beta}{\sigma^2} & \text{when } i = 2 \\ \frac{1}{\Phi^N\left(\frac{\mu+\theta}{\sigma}\right)} & \text{when } i = 3 \\ \exp -\frac{N}{2\sigma^2}(\mu-\theta)^2 & \text{when } i = 4 \\ \exp -\frac{1}{2\sigma^{*2}}(\log(\sigma)-\mu^*)^2 & \text{when } i = 5 \end{cases}. \tag{C.5}$$

We take their log

$$\log h_i = \begin{cases} -a\log\sigma & \text{when } i = 1 \\ -\frac{\beta}{\sigma^2} & \text{when } i = 2 \\ -N\log\Phi\left(\frac{\mu+\theta}{\sigma}\right) & \text{when } i = 3 \\ -\frac{N}{2\sigma^2}(\mu-\theta)^2 & \text{when } i = 4 \\ -\frac{1}{2\sigma^{*2}}(\log(\sigma)-\mu^*)^2 & \text{when } i = 5 \end{cases} \tag{C.6}$$

and compute the first two derivatives (relatively to μ and σ) for each of them to get

$$\frac{\partial}{\partial\sigma}(\log h_i) = \begin{cases} -\frac{a}{\sigma} & \text{when } i = 1 \\ 2\frac{\beta}{\sigma^3} & \text{when } i = 2 \\ N\varphi\frac{(\mu+\theta)}{\sigma^2} & \text{when } i = 3 \text{ (from (C.3))} \\ -\frac{1}{\sigma\sigma^{*2}}(\log(\sigma)-\mu^*) & \text{when } i = 5 \end{cases} \tag{C.7}$$

$$\frac{\partial^2}{\partial\sigma^2}(\log h_i) = \begin{cases} \frac{a}{\sigma^2} & \text{when } i = 1 \\ -6\frac{\beta}{\sigma^4} & \text{when } i = 2 \\ -N\left\{\varphi\left[-\frac{(\mu+\theta)}{\sigma}\frac{(\mu+\theta)^2}{\sigma^4}+\frac{2(\mu+\theta)}{\sigma^3}\right]-\varphi^2\frac{(\mu+\theta)^2}{\sigma^4}\right\} & \text{when } i = 3 \text{ from (C.4)} \\ = \frac{N\varphi\cdot(\mu+\theta)}{\sigma^3}\left\{\frac{(\mu+\theta)^2}{\sigma^2}+\varphi\cdot\frac{(\mu+\theta)}{\sigma}-2\right\} & \\ -\frac{1}{\sigma^{*2}}\frac{\left(\frac{1}{\sigma}\sigma-(\log(\sigma)-\mu^*)\right)}{\sigma^2} & \text{when } i = 5 \\ = \frac{\log(\sigma)-\mu^*-1}{\sigma^{*2}\sigma^2} & \end{cases} \tag{C.8}$$

$$\frac{\partial}{\partial\mu}(\log h_i) = \begin{cases} -N\frac{\phi}{\Phi\sigma} = -\frac{N\varphi}{\sigma} & \text{when } i = 3 \text{ (from (C.3))} \\ -\frac{N}{\sigma^2}(\mu-\theta) & \text{when } i = 4 \end{cases} \tag{C.9}$$

$$\frac{\partial^2}{\partial\mu^2}(\log h_i) = \begin{cases} -N\left\{\varphi\left[-\left(\frac{\mu+\theta}{\sigma}\right)\cdot\frac{1}{\sigma^2}+0\right]-\left[\frac{\phi}{\Phi\sigma}\right]^2\right\} & \text{when } i = 3 \text{ from (C.4)} \\ = \frac{N\varphi}{\sigma^2}\left[\frac{\mu+\theta}{\sigma}+\varphi\right] & \\ -\frac{N}{\sigma^2} & \text{when } i = 4 \end{cases} \tag{C.10}$$

C.4 Approximation of $\log(\Phi(g(z)))$ by Taylor series development

Let

$$f(z) = \log(\Phi(g(z))) \; . \tag{C.11}$$

A good approximation of $f(z)$ is obtained from the first three terms of its Taylor series development, that is

$$f(z) \approx f(0) + f'(0)z + \frac{f''(0)}{2} z^2 \; .$$

From (C.3), with $g(z) = z$, we easily obtain the first two derivatives of $f(z)$, respectively given by

$$f'(z) = \frac{\phi(z)}{\Phi(z)} \tag{C.12}$$

and

$$f''(z) = \frac{\phi'(z)\Phi(z) - \phi(z)\phi(z)}{\Phi^2(z)} \; ; \tag{C.13}$$

From (C.1), we obtain that $\phi'(0) = 0$; since $\phi(0) = \frac{1}{\sqrt{2\pi}}$ and $\Phi(0) = 1/2$ we easily get that

$$\begin{aligned} f'(0) &= \frac{1/\sqrt{2\pi}}{1/2} = \sqrt{\frac{2}{\pi}} \\ \text{and } f''(0) &= -\frac{\phi^2(0)}{\Phi^2(0)} = -[f'(0)]^2 = -\frac{2}{\pi} \; . \end{aligned}$$

Since $f(0) = \log(\Phi(0)) = \log(1/2) = -\log(2)$, we finally get the following second-degree approximation

$$f(z) \approx -\log(2) + \sqrt{\frac{2}{\pi}} z - \frac{1}{\pi} z^2 \; . \tag{C.14}$$

In some occasions, we may want to use a third-degree Taylor series development approximation. From (C.13), we have

$$f''(z) = \frac{\phi'(z)}{\Phi(z)} - \left[\frac{\phi(z)}{\Phi(z)}\right]^2$$

from which we easily get f's third derivative

$$f'''(z) = \frac{\phi''(z)\Phi(z) - \phi'(z)\phi(z)}{\Phi^2(z)} - 2\frac{\phi(z)}{\Phi(z)} \left\{ \frac{\phi'(z)\Phi(z) - \phi^2(z)}{\Phi^2(z)} \right\} \; ;$$

since

$$\begin{aligned} \phi'(z) &= -z\phi(z) \text{from (C.1)} && \text{(C.15)} \\ \text{we obtain } \phi''(z) &= -\phi(z) - z\phi'(z) \; . && \text{(C.16)} \end{aligned}$$

The Taylor series z^3-coefficient can now be easily calculated, as

$$\begin{aligned} f(0) &= \log(1/2) = -\log(2) \\ \phi(0) &= \frac{1}{\sqrt{2\pi}} \\ \Phi(0) &= \frac{1}{2} \\ \phi'(0) &= 0 \text{ (from (C.15))} \\ \text{and } \phi''(0) &= -\phi(0) = -\frac{1}{\sqrt{2\pi}} \\ \text{and finally } f'''(0) &= \frac{\phi''(0)}{\Phi(0)} + 2\left[\frac{\phi(0)}{\Phi(0)}\right]^3 = -\sqrt{\frac{2}{\pi}} + 2\sqrt{\frac{8}{\pi^3}} = \left(\frac{4}{\pi} - 1\right)\sqrt{\frac{2}{\pi}} \end{aligned}$$

from which we get the following third-degree approximation

$$\log(\Phi(z)) \approx -\log(2) + \sqrt{\frac{2}{\pi}}z - \frac{1}{\pi}z^2 + \frac{1}{6}\left(\frac{4}{\pi} - 1\right)\sqrt{\frac{2}{\pi}}z^3 \quad . \tag{C.17}$$

The second-degree approximation of $\log(\Phi(z))$ is good for $z \leq 1.5$ while its third-degree approximation is good for $z \leq 5.5$; figure below shows the very good fit for both versions in their respective domains.

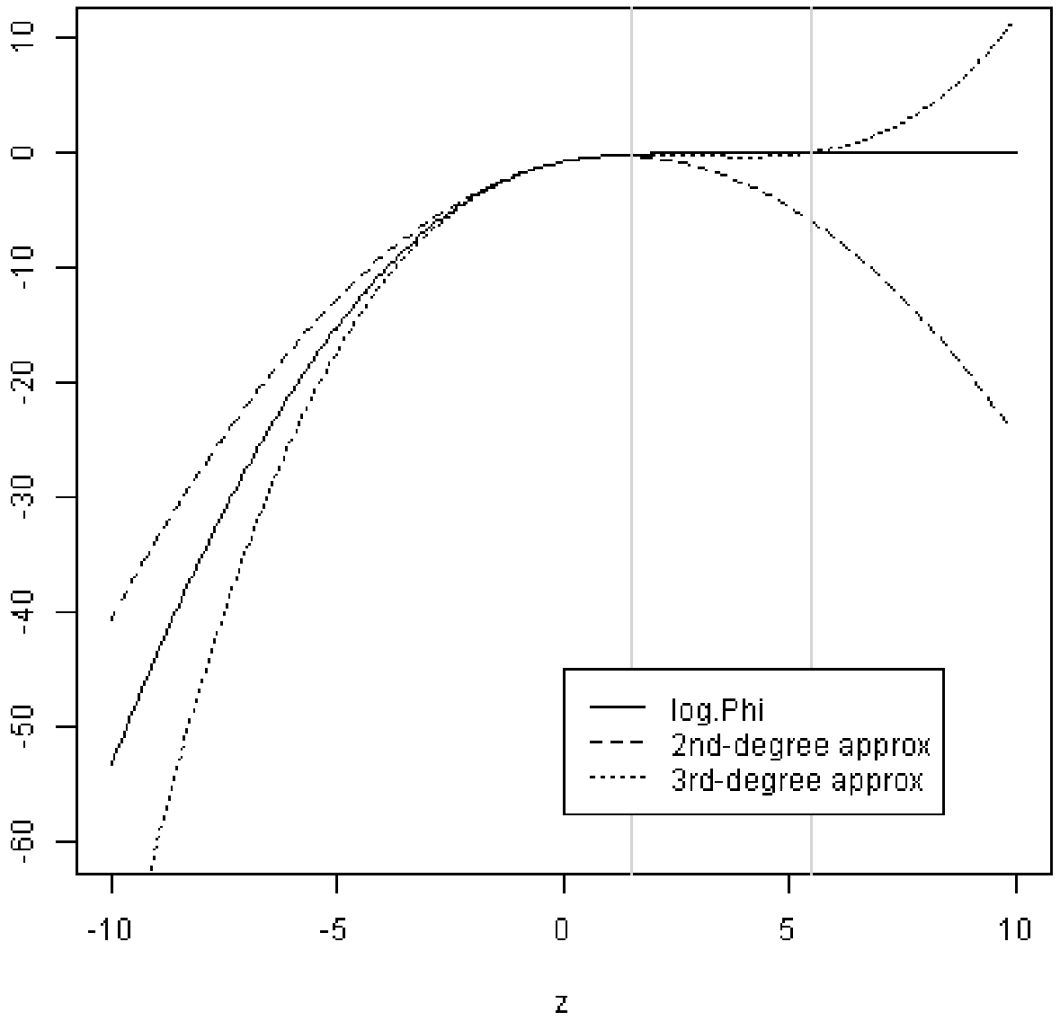


Figure 6: Second- and third-degree Taylor series approximation of $\log(\Phi(z))$.

C.5 Incorrect use of past data

The algorithm `SEG.informedvar` (section 3) including past data in the analysis was based on an improper prior (flat on $(0, \infty)$; while being technically correct, it was of limited interest, as no other algorithm was based on a uniform prior on σ. The current appendix documents that deprecated algorithm for historical reasons.

The likelihood of past data $\mathbf{p}$ (see equation 8)

$$l(\mathbf{p}|\mu,\sigma) = \underbrace{\frac{1}{\sigma^n}\exp\left\{-\frac{(n-1)s_p^2}{2\sigma^2}\right\}}_{g(\sigma^2)}\underbrace{\exp\left\{-\frac{n}{2\sigma^2}(\bar{p}-\mu)^2\right\}}_{l(\bar{p}|\mu,\sigma)}$$

is the product of the likelihood of $\bar{p}$ and a function $g(\sigma^2)$, which which we will approximate by a log-normal distribution with parameters (μ_p^*, σ_p^*) to be able to use the results of section 3. We approximate $g(\sigma^2)$ by a log-normal distribution for σ with parameters such that the means and 95% prior interval ranges match.

If we let $t = \sigma^2$, then the distribution function for t is easily obtained as

$$\begin{aligned} f(t) &\propto t^{-1/2}\frac{1}{(t^{1/2})^n}\exp\left\{-\frac{(n-1)s_p^2}{2t}\right\} \\ &= \frac{1}{t^{\frac{n+1}{2}}}\exp\left\{-\frac{(n-1)s_p^2}{2t}\right\} \end{aligned}$$

and hence $t = \sigma^2 \sim \text{InverseGamma}(\alpha = \frac{n-1}{2}, \beta = \frac{(n-1)s_p^2}{2})$.

If a random variable D has an Inverse-gamma distribution with parameters (α, β), then its mean is $E(D) = \log(\beta) - \psi(\alpha)$ where $\psi(\cdot)$ is the digamma function (see the inverse gamma distribution Wikipedia page).

Since σ^2 has an inverse gamma distribution, as shown above, we get

$$E(\log(\sigma^2)) = \log(\beta) - \psi(\alpha) \ .$$

The mean of its log-normal approximation μ_p^* is thus set to

$$\mu_p^* = \frac{\log(\beta) - \psi(\alpha)}{2} \ .$$

Let denote $\gamma_{\alpha,\beta;0.025}$ and $\gamma_{\alpha,\beta;0.975}$ the lower and upper 2.5 percentiles of a gamma distribution with parameters (α, β): then the lower and upper corresponding percentiles for $\log(\sigma^2)$ are given by $\log(\gamma^{-1}_{\alpha,\beta;0.975})$ and $\log(\gamma^{-1}_{\alpha,\beta;0.025})$ respectively, and hence the parameter σ_p^* is defined as

$$\sigma_p^* = \frac{1}{4\Phi^{-1}(0.975)}(\log(\gamma^{-1}_{\alpha,\beta;0.025}) - \log(\gamma^{-1}_{\alpha,\beta;0.975})) \ .$$

ANNEXE C: RESULTS SPECIFIC TO THE NORMAL MODEL

A. List of exposure metrics calculated for the normal distribution in the WebExpo project

Table C1: Exposure metrics calculated for the normal distribution in the WebExpo project

SEG analysis
Distributional parameter estimates (point estimate and credible intervals)
Arithmetic mean
Arithmetic standard deviation
Exceedance fraction of the OEL
Percentile of the exposure distribution (i.e., critical percentile, default 95%)
Decision on Exposure Acceptability (overexposure risk)
Probability that exceedance fraction ≥ exceedance threshold (default 5%)
Probability that critical percentile (default 95%) ≥ OEL
Probability that arithmetic mean ≥ OEL
Between-worker differences
Distributional parameter estimates (point estimate and credible intervals)
Group arithmetic mean
Within-worker arithmetic standard deviation
Between-worker arithmetic standard deviation
Within-worker correlation coefficient (rho)
Probability that rho is ≥ threshold (Prob.rho.overX)
R difference (R.diff)
Parameters quantifying the possibility that some workers are overexposed (probability of individual overexposure)
Proportion of workers with their individual critical percentile ≥ OEL (Prob.ind.overexpo.perc)
Proportion of workers with their individual arithmetic mean ≥ OEL (Prob.ind.overexpo.am)
Probability that the true value for Prob.ind.overexpo.perc is above a threshold (Prob.ind.overexpo.perc.overX, default 20%)
Probability that the true value for Prob.ind.overexpo.am is above a threshold (Prob.ind.overexpo.am.overX, default 20%)
For any individual worker : all metrics from the SEG analysis
Customizable parameters
Probability for credible intervals (default 90%)
Exceedance threshold (5%)
Critical percentile (default 95%)
Threshold for the within-worker correlation coefficient (default 0.2)
Coverage of the population for the R difference (default 80%)
Threshold for the probability of Individual overexposure (default 20%)

B. Interpretation of the Bayesian model outputs – SEG analysis

Figure C1 illustrates the data processing flow in the analysis for the normal distribution. For the normal model there is no pre-treatment of the observations prior to being processed by the Bayesian routines. Hence, users should specify the priors and measurement error in accordance with the scale of their quantity of interest. Default values for WebExpo were determined with noise exposure levels expressed in decibels (Appendix D).

The output metrics include arithmetic mean, arithmetic standard deviation, exceedance fraction of the OEL, any percentile of the distribution (default 95%), obtained from the equations below.

Arithmetic mean of the exposure distribution:

$$\boldsymbol{AM = \mu} \quad \textbf{(1)}$$

Arithmetic standard deviation of the exposure distribution:

$$\boldsymbol{ASD = \sigma} \quad \textbf{(2)}$$

Xth percentile of the exposure distribution:

$$\boldsymbol{PX = \mu + \Phi^{-1}(X) * \sigma} \quad \textbf{(3)}$$

where Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution.

Exceedance fraction of the OEL:

$$\boldsymbol{F(\%) = 100 * \left\{1 - \Phi\left(\frac{OEL - \mu}{\sigma}\right)\right\}} \quad \textbf{(4)}$$

where Φ is the cumulative distribution function of the standard normal distribution.

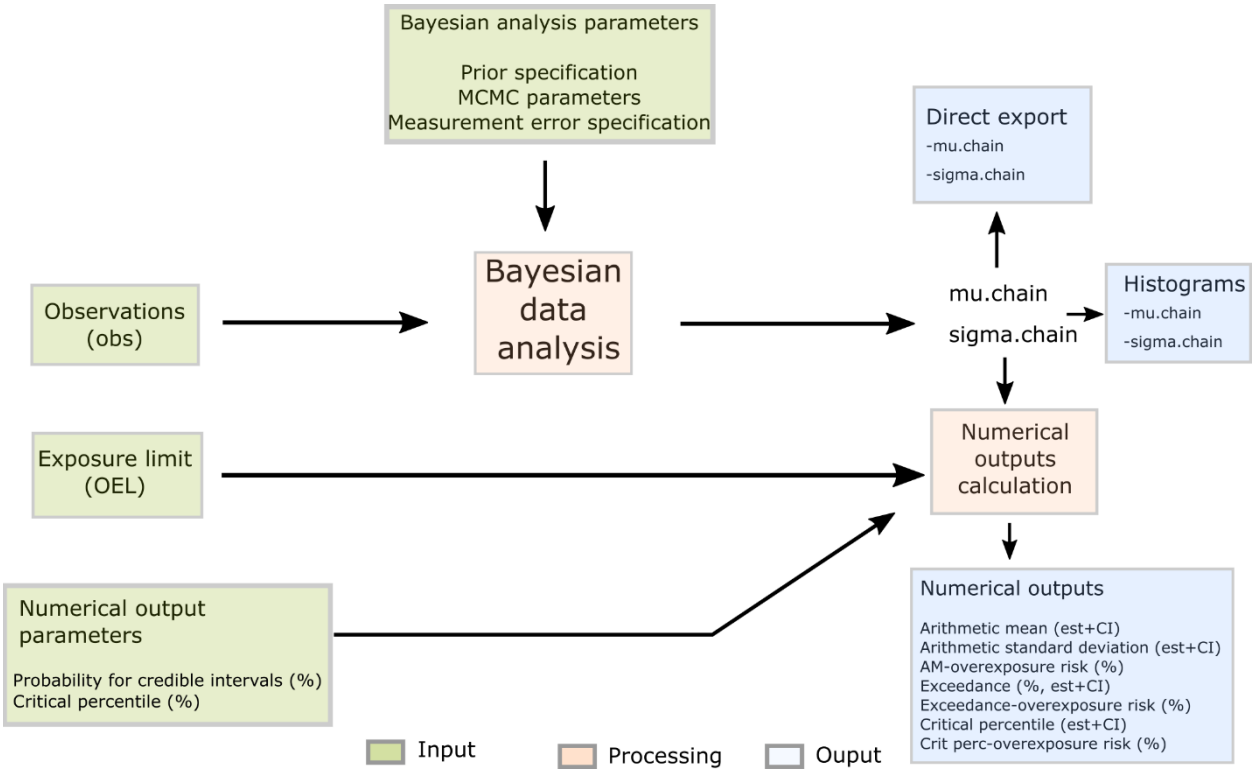


Figure C1 – Data processing flow for the SEG analyses – Normal distribution.

Example

We present below an example of the analysis of a sample of size 9 coming from a normal distribution with true AM=80 and true ASD=5 (considering an OEL of 85) [sample.C1 in Appendix E]. Table C2 presents the results of the analysis using the Bayesian model for the normal distribution and an uninformative prior using default parameters (see Appendix D). The very narrow width of the credible intervals shown in Table C2 illustrates the much lower variability of the normal distribution used in this example, which would be a plausible noise exposure distribution, when compared to the lognormal distribution used to model environmental variability of chemical exposures.

Table C2 – Exposure metrics point estimates and credible intervals for an example of Bayesian calculation for the normal model

Parameter	**Point estimates and 90% credible interval**
Arithmetic mean	78.6 [76.9 - 80.4]
Arithmetic standard deviation	3.01 [2.03 - 5.13]
Exceedance fraction (%)	1.72 [0.0517 - 13.7]
95^{th} percentile	83.6 [81.4 - 87.6]

C. Interpretation of the Bayesian model outputs – Between-worker analysis

Figure C2 illustrates the data processing flow in the analysis for the normal distribution. For the normal model (as for the normal SEG analysis model) there is no pre-treatment of the observations prior to the Bayesian routines. Hence, users should specify the priors and measurement error in accordance to the scale of their quantity of interest. Default values for WebExpo were determined with noise exposure levels expressed in decibels (Appendix D).

The following equations for normal exposure metrics were derived by adapting the lognormal equations above.

Group arithmetic mean:

$$\boldsymbol{AM_{group} = \mu_Y} \qquad \textbf{(5)}$$

Between-worker arithmetic standard deviation:

$$\boldsymbol{ASD_b = \sigma_b} \qquad \textbf{(6)}$$

Within-worker arithmetic standard deviation:

$$\boldsymbol{ASD_w = \sigma_w} \qquad \textbf{(7)}$$

Within-worker correlation coefficient:

$$\boldsymbol{\rho = \frac{\sigma_b^2}{\sigma_b^2+\sigma_w^2}} \qquad \textbf{(8)}$$

$R_{X\%}$ difference: Relative distance, expressed as a % of the group arithmetic mean, containing the middle X% of the distribution of either worker specific arithmetic means or any percentile. This proposal represents an attempt at expressing heterogeneity in a similar way as the R ratio for the lognormal model, adapted to the normal scale. Our proposed default value for X is 80%.

$$\boldsymbol{Rdiff_{X\%} = \frac{100 * \left(2 * \Phi^{-1}\left(\frac{1+X}{2}\right) * \sigma_b\right)}{AM_{group}}} \qquad \textbf{(9)}$$

Probability that a single random worker would have his own arithmetic mean above the OEL:

$$\boldsymbol{P_{ind}^{MA}(\%) = 100 * \left\{1 - \Phi\left(\frac{\mathrm{VLEP} - \mu_Y}{\sigma_b}\right)\right\}} \qquad \textbf{(10)}$$

Probability that a single random worker would have his own X^{th} percentile above the OEL (this is equivalent to the probability that a single random worker would have his own exceedance of the OEL above (100-X)%:

$$\boldsymbol{P_{ind}^{PX}(\%) = 100 * \left\{1 - \Phi\left(\frac{\mathrm{OEL} - \left(\mu_Y + \Phi^{-1}(X) * \sigma_w\right)}{\sigma_b}\right)\right\}} \qquad \textbf{(11)}$$

In addition to the above, it is also possible to obtain metrics specific to any individual exposure distribution. Hence by definition the exposure distribution for worker *i* is defined by:

Arithmetic mean of the exposure distribution:

$$\boldsymbol{AM = \mu_Y + b_i} \quad \textbf{(12)}$$

Arithmetic standard deviation of the exposure distribution:

$$\boldsymbol{ASD = \sigma_w} \quad \textbf{(13)}$$

X^{th} percentile of the exposure distribution:

$$\boldsymbol{PX = \mu_Y + b_i + \Phi^{-1}(X) * \sigma_w} \quad \textbf{(14)}$$

where Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution.

Exceedance fraction of the OEL:

$$\boldsymbol{F(\%) = 100 * \left\{1 - \Phi\left(\frac{OEL - \mu_Y - b_i}{\sigma_w}\right)\right\}} \quad \textbf{(15)}$$

where Φ is the cumulative distribution function of the standard normal distribution.

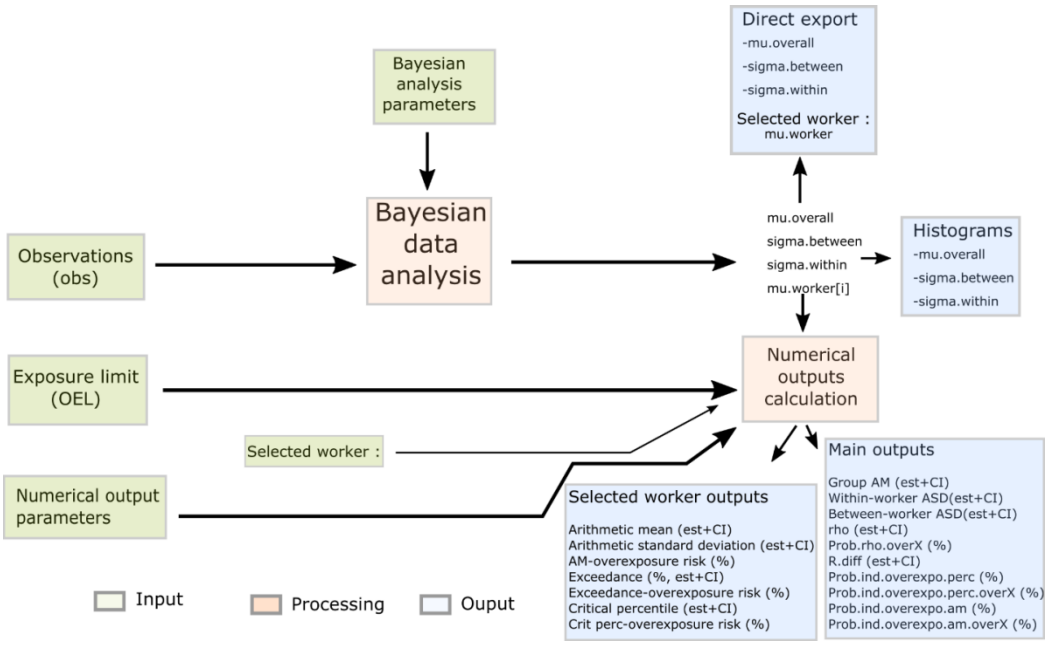


Figure C2 – Data processing flow for the between-worker difference analyses – Normal distribution.

Example

We present an example of the analysis of a sample of size 100 (ten repeats with ten workers) coming from a normal distribution with true AM=80 and true group ASD=5 [sample.C2 in Appendix E]. Without empirical information on realistic values for within-worker correlation, we used the median value found for chemicals in the Kromhout *et al.* database: rho=0.22. Table 11 presents the results of the analysis. As observed for the SEG analysis in 4.2.1.6, credible intervals shown in Table C3 are narrower than those for the lognormal case. For this example we assume no measurement error and we run the calculations with the [Between worker differences.informedvar] model implemented in R + RJAGS (see 4.3).

Table C3: Exposure metrics point estimates and credible intervals for an example of Bayesian calculation for the normal model (between-worker difference analyses)

Parameter	Low within-worker correlation
Arithmetic mean (90% CrI)	80.8 [78.8 - 82.7]
Between-worker arithmetic standard deviation (90% CrI)	3.17 [2.02 - 5.16]
Within-worker arithmetic standard deviation (90% CrI)	4.37 [3.88 - 4.96]
Within-worker correlation (rho) (90% CrI)	0.345 [0.166 - 0.591]
Probability that rho>0.2	90%
R.difference (90% CrI)	10 [6.4 - 16.4]
Probability of individual overexposure (95^{th} percentile) in % (90% CrI)	82.7 [59.1 - 96.8]
Chances that the above probability is >20%	100%
Probability of individual overexposure (95^{th} percentile) in % (90% CrI)	8.81 [1.01 - 29]
Chances that the above probability is >20%	15%

ANNEXE D:
INPUT PARAMETERS FOR THE BAYESIAN MODELS IN WEBEXPO

General data input instructions for all models

While in theory the Bayesian models can take as low as zero observation as input, in which case the posterior distributions will simply replicate the priors, we propose a "reasonable" minimal input as 3 uncensored results, with uncensored results representing at least 30% of the total sample.

For the lognormal models, the data should be strictly positive and between OEL/1000 and 1000*OEL. While being informal technical limits, these bounds are compatible with the proposed default parameters presented below for the Bayesian functions

For the Normal models, the data should be positive and away from zero due to the measurement error model. Based on noise data expressed in decibels, we propose that values should be between 40 and 140.

Uncensored observations should be entered as numbers, left censored observations as <X, right censored observations as >X, and interval censored observations as [X1-X2].

A. SEG.informedvar (including past.data model)

Parameter	Default recommendation	Reasonable range
n.chain[(A)]	1	1-5
n.iter	25 000	10000-100000
n.burnin	5000	100-5000
n.thin[(B)]	1	1
mu.lower	Lognormal : -20 Normal : 40	Lognormal : [-100 ; -0.5] and < min(observations) Normal (dB) : [20-85] et < min(observations)
mu.upper	Lognormal : 20 Normal (dB) : 125	Lognormal : [0.5 ; 100] and > max(observations) Normal (dB) : [85-140] et > max(observations)
log.sigma.mu	Lognormal : -0.1744 (GM=0.84) Normal : 1.098612 (GM=3)	Lognormal GM for the lognormal distribution of log(GSD) values is between 0.405 and 1.609, corresponds to log.sigma.mu between -0.90 and 0.48 Normal GM for the lognormal distribution of sigma is between 0.5 and 10, corresponds to log.sigma.mu between -0.69 and 2.30

Parameter	Default recommendation	Reasonable range
log.sigma.prec	Lognormal : 2.5523 Normal : 1.191059	Lognormal GSD for the lognormal distribution of log(GSD) values is between 1.5 and 5, corresponds to log.sigma.prec between 0.40 and 6.0 Normal GSD for the lognormal distribution of sigma is between 1.5 and 5, corresponds to log.sigma.prec between 0.40 and 6.0
init.mu	Lognormal : log(0.3) Normal : 85	Lognormal Between OEL/1000 and 100*OEL (init.mu between -6.908 and 4.605) Normal Between 50 and 120dB (init.mu between 50 and 120)
init.sigma	Lognormal : log(2.5) Normal : 3	Lognormal GSD between 1.5 and 10 (init.sigma between 0.405 and 2.303) Normal Sigma between 0.5 and 10 (init.sigma between 0.5 and 10)
past.data (mean)	N.A.	Lognormal Between OEL/1000 and 100*OEL (mean between -6.908 and 4.605) Normal Between 50 and 120dB (mean between 50 and 120)
past.data (sd)	N.A.	Lognormal GSD between 1.5 and 10 (sd between 0.405 and 2.303) Normal Sigma between 0.5 and 10 (sd between 0.5 and 10)
past.data (n)	N.A.	1-1000
me.sd.range	N.A.	Lognormal 0.001-1 (meaning between OEL/1000 and OEL) Normal 0.1-10
cv.range	N.A.	0.01-2 (meaning between 1 and 200%)

(A) While the number of MCMC chains is a parameter in the R code functions created by the McGill team, the relative simplicity of the models does not warrant multiple chains. Therefore in all other functions and in all other algorithms, this parameter is fixed to 1.

(B) Similar to note A, the thinning factor is set to 1 in all algorithms except for the R functions.

B. SEG.uninformative (restricted to parameters not described in A)

Parameter	Default recommendation	Reasonable range
sd.range	Lognormal GSD between 1.1 and 10 sd.range=[0.095-2.3] Normal ASD between 0.1 and 20 db sd.range=[0.1-20]	Lognormal GSD between 1.01 and 20 sd.range included in [0.01-3] Normal ASD between 0.1 and 100dB sd.range included in [0.1-100]
Init.sd	Lognormal : log(2.5) Normal : 3	Lognormal GSD between 1.5 and 10 (init.sd between 0.405 and 2.303) Normal Sigma between 0.5 and 10 (init.sd between 0.5 and 10)

C. SEG.riskband (restricted to parameters not described in A)

Parameter	Default recommendation	Reasonable range
A **(break points defining the bands)**	Lognormal (AIHA bands) 0.01/0.1/0.5/1 Normal 70/80/85/90	Increasing values from 0.001 to 100.
target_perc **(percentile of the distribution used to define the prior probabilities on the bands)**	95 (95th percentile)	1-99
region.prior.prob **(prior probabilities for each band)**	0.2/0.2/0.2/0.2/0.2	Any values between 0 and 1, provided the sum is 1
mu.lower.riskb	See mu.lower section A.	See mu.lower section A.
mu.upper.riskb	See mu.upper section A.	See mu.upper section A.
sigma.lower	See sd.range section B.	See sd.range section B.
sigma.upper	See sd.range section B.	See sd.range section B.

D. Between.worker differences

Constraints on observations are similar to the SEG functions. In addition, we recommend inputting at least 3 workers with at least 2 measurements each. Not all workers need to have repeated measurements. Note that the uncertainty surrounding worker specific exposure estimates will be directly related to the number of measurements for the worker of interest.

Parameter	Default recommendation	Reasonable range
n.iter	50 000	15000-200000
n.burnin	5000	500-10000
mu.overall.lower	Lognormal : -20 Normal : 40	Lognormal : [-100 ; -0.5] and < min(observations) Normal (dB) : [20-85] and < min(observations)
mu.overall.upper	Lognormal : 20 Normal (dB) : 125	Lognormal : [0.5 ; 100] and > max(observations) Normal (dB) : [85-140] and > max(observations)
log.sigma.between.mu	Lognormal : -0.8786 (GM=0.415) Normal : 1.098612 (GM=3)	Lognormal GM for the lognormal distribution of log(GSD) values is between 0.105 and 1.609, corresponds to log.sigma.between.mu between -2.25 and 0.48 Normal GM for the lognormal distribution of sigma is between 0.1 and 10, corresponds to log.sigma.between.mu between -2.30 and 2.30
log.sigma. between.prec	Lognormal : 1.634 Normal : 1.191059	Lognormal GSD for the lognormal distribution of log(GSD) values is between 1.5 and 5, corresponds to log.sigma. between.prec between 0.40 and 6.0 Normal GSD for the lognormal distribution of sigma is between 1.5 and 5, corresponds to log.sigma. between.prec between 0.40 and 6.0
log.sigma.within.mu	Lognormal : -0.4106 (GM=0.415) Normal : 1.098612 (GM=3)	Lognormal GM for the lognormal distribution of log(GSD) values is between 0.405 and 1.609, corresponds to log.sigma.mu between -0.90 and 0.48 Normal GM for the lognormal distribution of sigma is between 0.5 and 10, corresponds to log.sigma.within.mu between -0.69 and 2.30
log.sigma. within.prec	Lognormal : 1.9002 Normal : 1.191059	Lognormal GSD for the lognormal distribution of log(GSD) values is between 1.5 and 5, corresponds to log.sigma. within.prec between 0.40 and 6.0 Normal GSD for the lognormal distribution of sigma is between 1.5 and 5, corresponds to log.sigma. within.prec between 0.40 and 6.0

Parameter	Default recommendation	Reasonable range
init.mu.overall	Lognormal : log(0.3) Normal : 85	Lognormal Between OEL/1000 and 100*OEL (init.mu.overall between -6.908 and 4.605) Normal Between 50 and 120dB (init.mu.overall between 50 and 120)
init.sigma.between	Lognormal : log(2.5)=0.916 Normal : 3	Lognormal GSD between 1.1 and 10 (init.sigma.between between 0.095 and 2.303) Normal Sigma between 0.5 and 10 (init.sigma.between between 0.5 and 10)
init.sigma.within	Lognormal : log(2.5)=0.916 Normal : 3	Lognormal GSD between 1.1 and 10 (init.sigma. within between 0.095 and 2.303) Normal Sigma between 0.5 and 10 (init.sigma. within between 0.5 and 10)
Sigma.between.range	Lognormal GSD between 1.00 and 10 Sigma.between.range=[0-2.3] Normal ASD between 0 and 20 db Sigma.between.range=[0-20]	Lognormal GSD between 1.0 and 20 Sigma.between.range included in [0.00-3] Normal ASD between 0 and 100db Sigma.between.range included in [0-100]
Sigma.within.range	Lognormal GSD between 1.1 and 10 Sigma.within.range=[0.095-2.3] Normal ASD between 0.1 and 20 db Sigma.within.range=[0.1-20]	Lognormal GSD between 1.01 and 20 Sigma.within.range included in [0.01-3] Normal ASD between 0.1 and 100db Sigma.within.range included in [0.1-100]
me.sd.range	N.A.	Normal 0.1-10 Lognormal 0.001-1 (meaning between OEL/1000 and OEL)
cv.range	N.A.	0.01-2 (meaning between 1% and 200%)

ANNEXE E:
SAMPLES USED FOR THE NUMERICAL EXAMPLES

Sample 1: SEG analysis - main example

24.7	64.1	13.8	43.7	19.9	133	32.1	15	53.7

Sample 2: SEG analysis - measurement error example

96.6	38.3	80.8	15.1	34	73.4	14.5	64.8	27.4	48.7
43.3	43.4	57.8	94.9	44.1	44.3	62.9	117	51.6	64.7
50.1	74.7	221	46.8	84.3	93.4	126	46.9	29.5	73.8
66.9	61.3	30.2	101	22.6	191	29.3	68	114	33.7
52.5	118	49.7	60.4	36.6	55.9	31.9	84.3	75.8	39.5
28.3	56.5	44.2	48	36.6	70	37	72	48	66.1
72.4	80.9	69.1	162	67.3	75.2	40.5	25.6	44	120
56.3	42.9	6.63	24.9	40.9	81	97.2	74.7	79.6	48.8
75.3	54.8	66.5	71.3	28.7	87.5	51.9	19.6	60.8	45.9
46.9	84.8	120	103	36.7	92.7	32.8	73.8	214	65.3

Sample 3: Between-worker differences – low within-worker correlation

worker-1	worker-2	worker-3	worker-4	worker-5	worker-6	worker-7	worker-8	worker-9	worker-10
185	4.79	8.85	16.4	14.7	37.9	22	69.9	28.1	113
34.8	23	31.7	6.91	59.6	96.9	44.8	30.5	7.49	7.68
16.7	7.54	15.8	87.4	15	40.8	37.5	33.4	16	85.6
12.4	62.3	89.6	20	21.8	106	16.6	53	23	196
18.6	8.55	164	16.8	20.6	21.7	30.7	70.7	99.9	35
47.4	9.28	40.5	7.12	96.1	25.8	7.07	78.3	12	17.6
52.6	43.6	47.6	6.99	16.8	51.3	7.18	18	11.8	60.7
15.3	94.2	75.5	16.4	15.8	23	80.9	45.2	57.4	15.5
27.6	44.6	10.7	12.6	8.02	18.9	44.5	51.4	8.79	34.3
26.3	66.6	62.3	63.9	26.7	20.2	135	33.7	24	12.1

Sample 4: Between-worker differences – high within-worker correlation

worker-1	worker-2	worker-3	worker-4	worker-5	worker-6	worker-7	worker-8	worker-9	worker-10
66.8	14.2	186	23.5	43.8	41	6.56	9.21	19.6	78.7
46	53.9	84.6	16.2	31.1	11.4	9.5	9.42	14.3	28.2
61.1	21.8	94.4	40.2	13.1	4.44	6.97	28.7	22.8	41.3
54.6	47.8	218	130	24.1	12.9	5.92	72.9	35.1	14.4
31.7	48.8	189	42.2	27.7	22.7	2.42	35.6	28.9	72.9
74.3	76.5	130	25.7	23.9	20.5	14	17.2	36.9	10.2
60.9	41.3	107	35.4	40.2	12.6	12.3	20.2	13	16.2
53.4	20.4	80.6	40.8	60.3	8.35	3.07	13.4	13.3	15.8
38.9	31.9	288	109	29.8	13.6	7.01	10.5	13.6	42.2
27.5	31.1	173	40.9	37.2	28.1	6.49	26.3	37	61

Sample 5: Between-worker differences – realistic sample size

worker-1	worker-1	worker-1	worker-1	worker-2	worker-2
31	60.1	133	27.1	61.1	5.27
worker-2	worker-2	worker-3	worker-3	worker-3	worker-3
30.4	31.7	20.5	16.5	15.5	71.5

Sample 6: Variability of results across calculation platforms

<25.7	17.1	168	85.3	66.4	<49.8	33.2	<24.4	38.3

Sample C1: SEG analysis – normal sample

81	79.5	80.7	78.1	80.1	74.8	74.8	79.8	79.8

Sample C2: Between-worker differences – normal sample

worker-1	worker-2	worker-3	worker-4	worker-5	worker-6	worker-7	worker-8	worker-9	worker-10
76.2	70.6	79.2	79.1	85.3	77.8	79.1	80	80	89.1
82.3	78.7	77.7	77.6	92.2	89	80.7	76.6	81.2	85.4
81.7	77.6	73.5	81.2	75.8	81.9	85.8	84.6	73.8	81.8
73.7	76.9	78.9	82.6	84.1	80.4	84.8	77.1	80.7	88.1
79.4	79.5	81.6	81.6	76.1	88.5	88.5	81.5	76.9	86.4
79.1	84.8	83.1	82.4	84.6	87	82.6	77.4	77.5	81.6
80.2	77.6	85.1	76.9	78.9	85	78.6	73.5	74.6	86.8
71	65.5	84.2	87.6	75.8	88.1	90.1	82.2	70.6	81.4
86.9	74.1	79.8	80.4	89	81.3	82.9	74.4	82.3	86.7
75.6	69.9	84.1	79.7	87.1	90.6	83	77.6	66.4	83.6