Respirable Crystalline Silica in the Manitoba Construction Sector: Advancing Knowledge to Reduce Exposure

H Davies[1], M Gorman-Ng[2], J Lavoué[3]

[1] School of Population and Public Health, University of British Columbia
[2] BC Construction Safety Alliance
[3] Department of Environmental and Occupational Health, École de santé publique, Université de Montréal

Acknowledgements

This work was supported by a grant from the Research and Workplace Innovation Program of The Workers Compensation Board of Manitoba.

We wish to thank the members of the project advisory committee, Chris Hooter, IUPAT District Council 17 and Manitoba Building Trades; Michael Boileau, SAFE Work Manitoba; Derek Pott, Construction Safety Association of Manitoba; Phil McDaniel, Manitoba Heavy Construction Association.

We also wish to thank the many companies in the Manitoba construction sector who agreed to participate, as well as their employees who gave generously of their time, and shared their knowledge.

Executive Summary

Silica is one of the most common minerals on earth. It occurs in different forms; respirable *crystalline* silica (RCS; i.e., quartz particles that are less than 4 microns in diameter) is a confirmed cause of lung cancer and is also linked to the serious fibrotic lung disease silicosis. In Canada 570 lung cancers per year are caused by occupational RCS exposure, with 56% of these occurring in construction. The very low occupational exposure limit (OEL) for RCS reflects its high toxicity.

It has been estimated that there are 429,000 workers in Canada who are exposed to RCS and the largest fraction of workers exposed (62%) is in the construction industry. In Manitoba (MB), over 47,000 people are employed in the construction sector. Crystalline silica is present in many common construction materials and construction work involves mechanical tasks that release fine particles of RCS dust to the air. There is a demonstrated need for construction employers to reduce exposures to RCS. In order for employers to keep their workers' exposures below the OEL they need to be able to *quantify* exposure. This allows them to compare their exposure levels to the OEL as well as select *appropriate* control measures and personal protective equipment. However, construction worksites are highly dynamic, the physical environment is constantly changing and different work and tasks occur in different time and space patterns and contractors frequently move between worksites. This can make it difficult for employers to measure and understand the exposure levels that might be present in their work places.

To assist construction employers with exposure estimation and with risk assessment, the authors had previously developed a database of construction related RCS exposure data, and models for estimating RCS exposures in construction for risk assessment purposes. These formed the basis of the web-based tool that uses objective, quantitative exposure data that has been previously obtained and resides in a database in the preparation of "exposure control plans" (ECP's). The ECP clearly conveys a quantitative assessment of risk, and estimates of the risk reduction afforded by different control measures, without having to undertake expensive and time-consuming field sampling for each new job. The tool was originally developed for the British Columbia (BC) construction industry. There has been considerable interest in the tool from outside of BC, and the developer of the on-line tool (The BC Construction Safety Alliance) has endeavoured to make the on-line tool more widely available for other jurisdictions in Canada.

This project was designed to reduce the risk of RCS-related disease in Manitoba by (i) increasing our understanding of RCS exposure through exposure monitoring on MB construction worksites; (ii) by improving the on-line tool through improvements to its design and readying it for use in Manitoba, and (iii) through KT activities to raise awareness and inform targeted groups about the RCS hazard and its control.

We undertook a comprehensive assessment of RCS exposure in Manitoba construction worksites between June 2018 and August 2020. A total of 121 measurements were made in 14 different company sites. Seventy-four measurements (61%) were obtained in urban Winnipeg, and the rest in smaller municipalities of Manitoba. Supplementary data was collected for each measurement to allow for statistical analysis of the determinants of exposure to RCS, in particular to examine regional differences in exposure levels.

Overall RCS had a geometric mean (GM) of 0.033 mg/m^3 which is higher than the regulatory OEL of 0.025 mg/m^3. Renovation work tended to have higher exposure levels and other types of construction project (GM = 0.084 mg/m^3) and demolition tasks tended to have the highest exposures (GM = 0.61 mg/m^3) followed by breaking and grinding. Working with cement resulted in the highest mean measured exposures (GM=0.083 mg/m^3). Working in restricted spaces (versus indoor/outdoor) was associated with the highest exposure levels (GM = 1.132 mg/m^3). Notably, use of individual controls mounted on tools did not appear to greatly reduce exposure levels. However, when both exhaust ventilation and wetting were used together exposure levels were reduced by about half.

Determinants of exposure modeling showed that RCS exposure levels in Manitoba in this study were approximately half that of those found in BC, but double those found in Alberta (AB). These differences may be attributable to differences in work practices, but also source material composition, and regulatory environment in the different provinces (noting workplace health and safety is mostly provincially regulated).

The current statistical prediction model was improved in two areas. First, new data from BC, AB and MN as well as from an updated literature review were added to the model's database. Second, updated statistical programming improved the estimation of uncertainty around the model's predictions. The prediction model was also improved to include two risk analysis frameworks to communicate the exposure risks based on the above model prediction: (i) the risk assessment framework from British Columbia was implemented, where the main risk metric is the geometric mean (GM) of the distribution of workers' exposures, and risk is considered controlled if the GM is below the OEL.; and (ii) the risk analysis framework from the American Industrial Hygiene Association (AIHA), where the main risk metric is the 95th percentile (P95) of the distribution of workers' exposures and risk is considered controlled if P95 is below the OEL. The statistical model was also improved to combine information from the silica prediction model with a new data set of exposure measurements which might have been collected for a particular situation. This allows the model to use the large amount of archived data held in the database (maybe hundreds of data points) efficiently while also taking into consideration the current exposure measurements (maybe just a handful) specific to the risk assessment at hand. Finally, these advances were combined into a new, improved user interface. A prototype can be viewed at https://silica.expostats.ca/

The authors make recommendations with respect to:

- approaches to reducing RCS exposure hazards. Notably use of LEV and engineering controls did not demonstrate dramatic reductions in exposure levels. This suggests that the mere presence of controls may not be sufficient; these interventions must be implemented effectively. Also, that quantitative risk assessments can be used to plan administrative controls (such as job rotation) to reduce long-term average exposures
- improving hazard awareness through knowledge translation. Study data can be used for training and education of (i) silica health hazard; (ii) hazardous exposures to RCS; (iii) controlling RCS exposure. Target audiences may include employees, employers, apprenticeship trainers, health and safety regulatory officers, claims adjudicators and others in compensation and prevention.
- upgrades to the BC Construction Safety Alliances "Silica Control Tool™", as well as the new end-user interfaces for risk assessment.

Future work should include additional sharing of results from this project with various stakeholders in Manitoba and the scientific community through peer-reviewed publications, and seeking funding and partners to integrate silica control tool improvements into the BCCSA tool.

Table of Contents

List of Figures and Tables

List of Acronyms and Abbreviations

ALARA	As Low As Reasonably Achievable
BCCSA	BC Construction Safety Alliance
COPD	Chronic Obstructive Pulmonary Disease
CSAM	Construction Safety Association of Manitoba
CSP	Common silica-producing process
ECP	Exposure Control Plan
IARC	International Agency for Research on Cancer
LOD	Limit of Detection
LPM	Liters per minute
OCRC	Occupational Cancer Research Center
OEL	Occupational Exposure Limit
OHS	Occupational Health and Safety
RCS	Respirable crystalline silica
SCT	(BC Construction Safety Alliance) Silica Control Tool
SME	Small and medium enterprise
TLV® - TWA	ACGIH Threshold limit value - Time-Weighted Average
UBC	University of British Columbia

1 Introduction & Overview

1.1 Silica Hazard

Silica is one of the most common minerals on earth. It occurs in different forms; respirable crystalline silica ("RCS", i.e., $\propto -$ quartz particles that are less than 4 microns in diameter) is a confirmed cause of lung cancer (IARC, 2012). RCS exposure is also linked to the serious fibrotic lung disease silicosis (Bang *et al.*, 2015), and has possible links to stomach and gastric cancers, COPD and autoimmune disease (NIOSH, 2002). Its very low occupational exposure limit (OEL, TLV®-TWA 0.025 mg/m^3) reflects its high toxicity. It has been known as a health concern for centuries but remains responsible for significant disease burdens in working populations, even in modern society with advanced OHS regulatory systems.

1.2 Silica and construction

There are an estimated 429,000 workers in Canada who are exposed to RCS (CAREX Canada, 2021). The largest fraction of workers exposed to RCS is in the construction industry, where 62% of workers are estimated to be exposed. In Manitoba, over 47,000 people are employed in the construction sector (Statistics Canada), suggesting almost 37,000 Manitobans might be at risk from RCS exposure. In addition, there are many thousands more in mining, agriculture and manufacturing jobs where RCS exposures are also likely. Previous work by the applicants (and colleagues) at the Occupational Cancer Research Center estimates that in Canada 570 lung cancers *per year* are caused by occupational RCS exposure, with 56% of these occurring in construction employees (OCRC, 2019).

Crystalline silica is present in common construction materials such as concrete, cement, brick, tiles, drywall, rock, sand and asphalt. Construction work involves mechanical tasks that generate fine RCS dust particles that enter the air (Beaudry *et al.*, 2013). A recent study of RCS exposure in Alberta found that overexposure to RCS is common in construction. Among workers involved in construction of new commercial buildings, seventy-seven percent of exposures were above the OEL. Overexposures were also common in other construction sectors, including demolition (forty percent above the OEL) and earth moving/road building (twenty five percent above the OEL; Radnoff *et al.*, 2014).

Employment in the construction industry in Manitoba is forecast to remain steady over the next decade, but with a turnover of approximately 20% (Build Force Canada, 2021)

1.3 Determinants of RCS Exposure

Difference between work task is a major contributor to the variability in construction RCS exposures (Beaudry *et al*, 2013). For example, tasks such as demolition, abrasive blasting, tuck-pointing using grinders, jackhammers, and drilling on materials like concrete are common tasks with high RCS exposures across different studies (Flanagan *et al.*, 2003, 2006; Rappaport *et al.*, 2002; Sauvé *et al.*, 2012). Exposure levels may also be expected to vary significantly due to differences within tasks, work techniques, control measured, project type, sampling duration, and other environmental factors (IARC, 2012; Radnoff *et al.*, 2015). The use of engineering controls such as local exhaust ventilation and wet dust suppression have been found to reduce exposure, while working indoors can increase RCS exposure (Rappaport *et al.*, 2002; Croteau *et al.*, 2004; Sauvé *et al.*, 2013).

1.3.1 Regional variability

Regional variability in RCS exposure levels might be anticipated, for example due to policy and regulatory differences, and particularly in Canada where OHS regulation (such as OEL setting) is a provincial responsibility. There may also be regional differences that occur due to the geologic differences in source materials such as gravel and sand. High percentages of quartz are commonly found in sedimentary rocks (i.e., sandstones) that are more typical of rock types in Alberta and Manitoba than BC (Atkinson & Atkinson, 1978; Carmichael, 1989; Heaney & Banfield, 1993; IARC,1996; Ross *et al.*, 1993; US Bureau of Mines, 1992). Further there may be regional differences related to overall "safety-culture" stemming from workforce demographics, training etc.

1.4 RCS Risk Assessment

Exposure monitoring is the gold standard for a comprehensive worksite risk assessment, but construction sites are complex and highly dynamic workplaces with high day-to-day and site-to-site variability, complicating exposure assessment. Many different work and tasks occur in different time and space patterns and sub-contractors move frequently between worksites. This can make it difficult for employers to measure interpret and understand RCS exposure levels that might be present in their workplaces.

In response to this challenge, the authors in collaboration with WorkSafeBC and the BC Construction Safety Alliance (BCCSA) developed a novel, on-line risk assessment tool for RCS exposure (the "Silica Control Tool™", SCT) for use across the construction sector (Table 1). The SCT provided a scientifically sound aid to the introduction of new regulation in BC that permitted the substitution of actual exposure measurement data with "objective air monitoring data", collected at "equivalent work operations." This allowance encouraged the introduction of quantitative risk assessment in a construction work environment where routine exposure monitoring would be very challenging for the employer.

Table 1: Benefits of the quantitative RCS risk assessment SCT for different stakeholder groups

Employee	Employer	Researcher/policymaker
Risk Awareness	Risk Awareness	Exposure data collection
Risk awareness	Identification of best practice	Education tool for regulators
Selection of appropriate control	Cost and time efficiency	Intervention tool for officers, demonstrate best practice
Clear ECP to follow	Standardized practice	Quicker uptake of best practices for an ALARA substance
	Comparison data (inter-site, inter-province, etc.)	

The SCT has a continually updatable database of RCS exposure measurements. The database was used to derive a predictive statistical model that can generate exposure risk estimates (Figure 1). The statistical model is embedded in an adaptive web-based application that can be run on common platforms. At the outset, the SCT was based on largely historical exposure data obtained from a database to which author JL contributed (Beaudry *et al*. 2013). Subsequently, the exposure database has been expanded with contemporary measurements representing conditions in British Columbia and Alberta workplaces, and more up-to-date worksite environments.

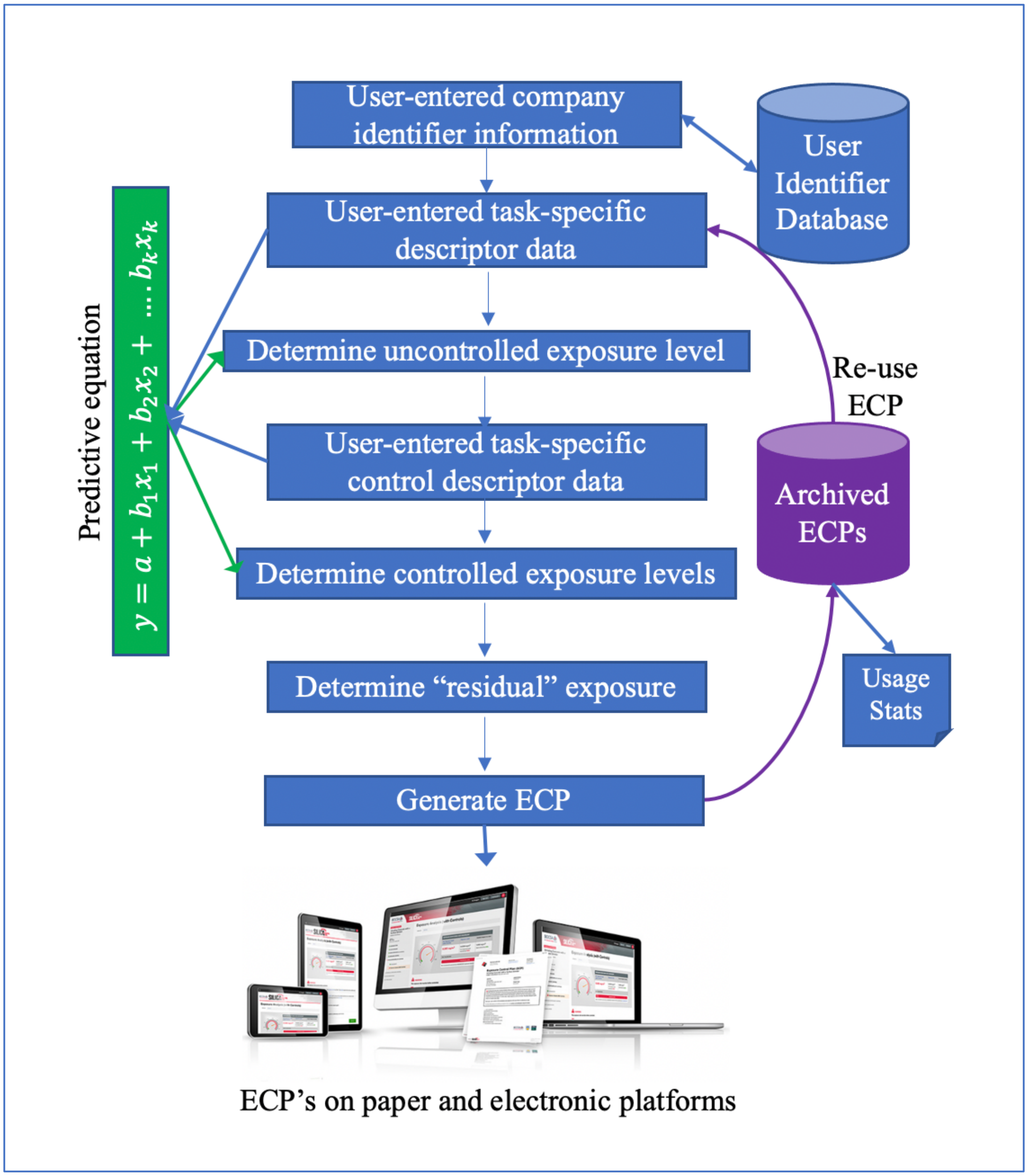

Figure 1: Silica Control Tool Components. The predictive equation (green) is generated from RCS data collected from published literature but also contemporary exposure data collected by the authors in BC, AB and now, Manitoba.

The earliest versions of the SCT were designed with "frequentist" statistical models that had limitations in how variability ("uncertainty") around an exposure estimate could be calculated and presented to the end-user. The current SCT (Davies & Gorman-Ng, 2020) estimates uncertainty by running the predictive equation many hundreds of times, using a technique called "bootstrapping", each time varying input parameters based on their underlying distribution. The estimated (predicted) exposure concentrations for a given work scenario is then presented to the end user as a single number (the point estimate from the predictive equation plus the 95% confidence interval from the bootstrap process).

There is scope for this approach to be modernized with Bayesian statistics. Exposure scientists have developed improvements in methodology for estimating risk using Bayesian statistics (a "non-frequentist" approach), and in improving how we communicate the complicated concepts around exposure risk (Lavoué, *et al*, 2018). Bayesian approaches permit calculation of simpler probabilistic risk estimates; such probabilistic estimates are considered to be easier to interpret than the estimates currently used in the SCT.

Furthermore, while the current SCT provides information from the historical database, users who have collected their own measurements do not have means to combine them with the tool's output. There are, however, established techniques (e.g., Bayesian approaches) that would allow us to extract the maximum information *simultaneously* from both historic and contemporary data. Bayesian approaches would permit end-users of the SCT (and other tools designed like it) to benefit from the data stored in the exposure database while combining it with their own company's exposure data, for example.

1.5 Study Objectives

There is a demonstrated need for construction employers to reduce exposures to respirable crystalline silica (RCS). In order for employers to keep their workers' exposures below the OEL they need to be able to quantify exposure. This allows them to compare their exposure levels to the OEL as well as select appropriate control measures and personal protective equipment (Kromhout, 2016). This poses a challenge for large companies with in-house occupational health and safety staff and an even greater challenge for small and medium sized enterprises (SMEs) who may not have in-house OHS expertise.

The project aimed to aid risk reduction for RCS-related disease in Manitoba and other provinces by increasing the understanding of RCS exposure through exposure monitoring on Manitoba construction worksites, by improving the online tool through improvements to its design, and through knowledge translation activities to raise awareness and inform targeted groups about the RCS hazard and its control.

Study Objectives:

a. Overall, to improve understanding of RCS hazard in MB construction industry
b. To characterize RCS exposure situation in MB construction, and compare to other Canadian regions
c. To improve RCS exposure data holdings vis-à-vis regional, contemporary and new exposure scenarios
d. Improve modeling and estimation methodology and evaluate, both generally, and with respect to use in Manitoba

2 Methodology

2.1 Ethical Review

Ethical approval of this study was provided by the UBC Behavioural Research Ethics Board, reference number H18-01017.

2.2 Exposure Monitoring

2.2.1 Recruitment

Companies were recruited through industry safety associations such as the Construction Safety Association of Manitoba (CSAM), and individual contacts. Individual employees were recruited on-site and selected based who was performing targeted work.

To focus our sampling efforts, "common silica -generating processes" (CSPs) were identified through focus groups of construction experts, and then reviewed in a survey of MB construction experts. For exposure sampling, we focused on these CSPs, including both controlled and uncontrolled exposure where possible (see full list of CSP's in Appendix 1).

2.2.2 Exposure sampling

Sauvé *et al* (2012) found that statistical models derived from task-based RCS exposure samples were more predictive of RCS exposure levels than models based on full-shift samples, so we aimed to collect task-based samples in most instances. Full-shift sampling was conducted when workers were conducting the same task for the entire day, and when we aimed to characterize the exposure associated with a job title (e.g., concrete plant operator) that was not directly involved in generating RCS. We defined "task" as including all activities related to the task from job set-up to final clean-up.

Sampling for RCS (e.g., NIOSH method 7500) typically involves the use of a cyclone at flow rates ranging from 1.7 to 2.5 liters per minute (LPM) to sample the respirable fraction of dust. However, when sampling durations are much shorter than eight hours, flow rates below 2 LPM may not collect sufficient quantities of crystalline silica to be detectable by chemical analysis (e.g., a typical analytical limit of mass detection is 0.005 mg). To address this, we used a "high-flow method" when sampling durations were short and/or ambient concentrations were expected to be low. The high-flow method we used was described by Stacey and Thorpe (HSE, 2010) and used a parallel particulate impactor (PPI) at a flow rate of 8 LPM (SKC Inc., Eighty-Four, PA, USA). Thirty-five mm PVC filters were used with both cyclone and PPI sampling heads. Approximately 10% percent of samples were designated as field blanks. We also collected comprehensive supplementary information on task, material, tool, work environment, use of exposure controls, and characteristics of the construction company and worksite (see Appendix 2) and took photographs to aid later interpretation of data.

All samples were analyzed by an AIHA Accredited industrial hygiene laboratory (Bureau Veritas North America, Novi, MI, USA). Samples were analyzed for cristobalite and quartz by x-ray powder diffraction (NIOSH method 7500). All samples were blank-corrected. Unless otherwise stated, RCS exposures combine quartz and cristobalite concentrations.

2.3 Data Analysis and Database Updates

2.3.1 Descriptive statistics

Data analysis of Manitoba exposure data was conducted using R version 4.1.0. Inspection of data distributions indicated that RCS concentrations were log-normally distributed so all analyses were performed on log-transformed data. RCS concentrations were summarized with geometric mean, geometric standard deviation, and minimums and maximums. Samples below the limit of detection (<LOD or ND) were substituted with LOD/2).

2.3.2 Exposure modeling

Multiple linear regression models were built after adding Alberta and BC data for regional comparison purposes. We selected variables for the regression model by first conducting bivariate analyses (chi-square and ANOVA tests). Correlations were explored to examine potential multicollinearity. Retained independent variables were included in a multivariable forward-stepwise linear regression model using the OLSRR package in R. These variables included province, task type, control, material, work environment, sampling duration, project type and activity sector. Many of these variables were previously identified by Sauvé *et al.* (2012) as determinants of RCS exposure in the construction industry and used to develop the original BCCSA Silica Control Tool (Davies and Gorman-Ng, 2020). The dependent variable was the log-transformed RCS concentration. Additional models restricting data to only task-based samples, excluding full-shift samples, were also conducted to test the robustness of the model. Task-based samples are taken to assess exposure associated with a particular task and may be more closely related to the "task type" variable, whereas full-shift samples may vary in the amount of time spent doing silica generating activities. Partial F-tests were conducted for all independent variables included in the final models.

2.4 Improvements to Exposure Risk Assessment Tools

2.4.1 Improvement of the current SCT prediction model

The current statistical prediction model was improved in two areas. First, new data from BC, AB and MN as well as from an updated literature review were added to the model's database. Second, meta-analysis and multi-model inference theory were used to improve the estimation of uncertainty around the model prediction. This procedure allows to obtain a model prediction in the form of a geometric mean of silica exposure with a standard error, which is both easier to manipulate and less computationally costly than the bootstrap procedure mentioned above.

2.4.2 Improvement of the risk metrics calculated from the SCT model

Two risk analysis frameworks were employed to communicate the exposure risks based on the above model prediction.

First, the risk assessment framework from British Columbia was implemented, where the main risk metric is the geometric mean (GM) of the distribution of workers' exposures (WorkSafeBC, 2021). Risk is considered controlled if the GM is below the OEL. The point estimate and standard error of the exposure model predicted GM described above is used to show the predicted geometric mean, a 95% upper confidence limit, as well as the probability that its true underlying value is above the OEL, i.e., the probability of non-compliance.

Second, the risk analysis framework from the American Industrial Hygiene Association (AIHA) was implemented, where the main risk metric is the 95th percentile (P95) of the distribution of workers' exposures (Jahn *et al.*, 2015). Risk is considered controlled if P95 is below the OEL. Furthermore, the AIHA proposes to further separate risk into management bands depending on the value of P95 compared to the OEL (see Figure 2).

	*SEG Exposure Risk Rating***	*Applicable Management/Controls*
	0 (<1% of OEL)	no action
	1 (<10% of OEL)	procedures and training, general hazard communication
	2 (10-50% of OEL)	+ chemical specific hazard communication, periodic exposure monitoring
	3 (50-100% of OEL)	+ required exposure monitoring, workplace inspections to verify work practice controls, medical surveillance, biological monitoring
	4+ (>100% of OEL, Multiples of OEL; e.g., based on respirator APFs)	+ implement hierarchy of controls, monitoring to validate respirator protection factor selection

Figure 2: AIHA Risk Management Bands (from Jahn, et al., 2015)

If we had access to the actual exposure distribution, i.e., all exposure levels for all workers in an exposure group over a stable period of time, we would know the true value of P95 and it would be straightforward to select the right management category. Unfortunately, we estimate

P95 from a very limited sample from this population, and therefore need statistical tools to evaluate the associated, and usually large, uncertainty. Therefore, the output created for this project provides a probability for each of the management bands: the probability that the true underlying P95 of exposure is in the band, summing to 100% across the 5 bands. The probability of the highest risk band (red band, 4+) is the probability of non-compliance, i.e., the probability that the 95th percentile of the distribution of exposures is above the OEL. The 95th percentile of a lognormal distribution depends on both the geometric mean (GM) the geometric standard deviation (GSD). The prediction model (see 2.4.2) provides point estimate and uncertainty for GM, but not for GSD. Therefore, we propose to use a typical workplace GSD value of 2.5, surrounded by uncertainty so that 90% of possible values are between 1.7 and 4.7 (Lavoué *et al.*, 2018). The point estimates and uncertainties for GM and GSD are then combined to estimate a point estimate for P95, a 95% upper confidence limit, and the probabilities associated with each AIHA category.

2.4.3 Bayesian model for combining the model prediction with a "new" silica dataset

Bayesian statistics are a field of statistics of increasing use in occupational hygiene (Banerjee *et al.*, 2014; Hewett *et al.*, 2006; Lavoué *et al.*, 2018; Quick *et al.*, 2017). In essence, each parameter that we want to estimate must be defined a priori as a probability distribution: For example, for the 95th percentile of the distribution of exposures, we could state that every value between a 1/100th of the OEL and 100 times the OEL is equally likely, which is a uniform distribution on the range "OEL/100-100*OEL". Then this distribution (called the "prior" distribution) is updated with the information collected (i.e., the measurements) through a statistical model (in this case it is assumed that exposure measurements follow a lognormal distribution). The result of the analysis is an updated probability distribution (called posterior), from which point estimates and confidence limits can be derived.

One considerable advantage of this framework relates to the flexibility in defining the prior distribution: our example demonstrated what is called an uninformative prior, i.e., it contained very little information on P95. In that case, the posterior distribution will mostly rely on the observations. However, it is possible to define an informative prior distribution, i.e., something is known of P95 in advance of collecting new data. This ability has been used in industrial hygiene for example using the output of mechanistic two-zone models to define the prior distribution of exposure parameters (Zhang *et al.*, 2009).

In this project, we set up a Bayesian model which allows us to combine information from the silica prediction model with a new data set of exposure measurements which might have been collected for a particular situation. The advantage of this combination is that both sources of information are limited on their own: the model is based on a lot of data, but provides estimates for generic scenarios and might not be entirely representative of any particular settings; on the other hand, measurement sets for any particular situation tend to be small, and though more relevant in theory, their small size renders direct extrapolation for risk assessment hazardous. The combined exposure estimate therefore results in an overall improved assessment.

The framework for building this Bayesian model was based on the WEBEXPO project (https://www.irsst.qc.ca/en/publications-tools/publication/i/101066/n/webexpo). The library of Bayesian models built within WEBEXPO was extended to include a type prior distribution defined by the output of a statistical model such as the SCT model.

2.4.4 Developing improved user interface

The prediction model (2.4.2) and Bayesian model (2.4.3) were combined into an online prototype for an improved user interface for risk assessment. The aim was first to allow users to select a prediction scenario closest to their exposure situation of interest, and see what this scenario entailed in terms of risk according to the silica database. In a second step, the user could enter measurement data relevant to his situation, and see what these data alone entailed in terms of risk. Finally, the user would be shown the result of the combined analysis, merging the information from both sources of information. A "recap" page would show the three sets of information on the same page for comparison purpose.

3 Results

3.1 Respirable Crystalline Silica Exposure Assessment

Fourteen Manitoba construction companies participated in this part of the study. Due to untypical spring and fall weather and the onset of the global Covid-19 pandemic, the worksite monitoring phase was extended and took place over two construction seasons, between June 2018 and August 2020. The majority of measurements (55%) were obtained during summer months between June and August, and the remainder sampled in the months of September and October.

A total of 121 measurements were made. Seventy-four measurements (61%) were obtained in urban Winnipeg, and the rest in smaller municipalities of Manitoba. Our exposure measurements were short-term, task-based samples of average duration of 55 minutes (range 10 minutes to 130 minutes).

Table 2 shows the summary statistics for RCS exposure concentrations. Overall RCS had a geometric mean of 0.033 mg/m^3 (arithmetic mean 0.26 mg/m^3). The GSD for RCS was 7.0, with the minimum and maximum concentration being <0.0003 mg/m^3 and 9.3 mg/m^3 respectively. Table 2 also shows recent results from recent Alberta and British Columbia sampling campaigns (also by the authors) for comparison purposes.

Table 2: Descriptive Statistics for Respirable Crystalline Silica Exposure by Province

Province	N (%)	%ND (<LOD)	GM (mg/m^3)	GSD	MIN (mg/m^3)	MAX (mg/m^3)
Manitoba	121 (32)	37%	0.033	7.0	< 0.001	9.3
Alberta	129 (35)	29%	0.060	9.3	< 0.004	8.2
British Columbia	123 (33)	19%	0.044	4.3	< 0.005	3.5

Table 3 shows the measured exposure by company. The number of samples varied by company, from two (site #2) to 41 (site #7). Site #1 had the highest average concentration of

RCS (N=10 samples). Site 14 had the lowest average concentration. Nine of the fourteen companies participating had GMs over the Manitoba OEL of 0.025 mg/m³).

Table 3: Company data (highest to lowest Geometric Mean exposure)

ID	N	GM mg/m³	GSD	MIN mg/m³	MAX mg/m³
1	10	0.489	3.8	0.046	9.300
2	2	0.230	1.5	0.170	0.310
3	4	0.207	8.0	0.016	1.700
4	4	0.083	2.1	0.035	0.160
5	5	0.064	2.6	<0.041	0.200
6	3	0.051	35.8	<0.013	3.200
7	41	0.040	6.6	<0.001	2.900
8	4	0.031	6.9	<0.001	0.240
9	10	0.030	3.4	0.009	0.350
10	5	0.013	11.5	<0.007	1.000
11	2	<0.021	1.0	<0.021	<0.021
12	5	<0.026	1.7	<0.007	<0.026
13	22	0.007	2.9	<0.005	0.250
14	4	< 0.013	1.1	<0.011	<0.013
121	121	0.0331	7.0	< 0.001	9.3

Most projects (Table 4) measured were renovation (45%) followed by new construction (40%). Both new construction and other not specified projects had similar levels of average RCS concentration (GM 0.0163 and 0.0166 mg/m³ respectively). Minimum RCS exposure occurred in demolition while maximum exposure occurred in renovation projects.

Table 4: RCS exposure concentrations by project type (by descending GM)

	N	GM (mg/m³)	GSD	MIN (mg/m³)	MAX (mg/m³)
Renovation	54	0.084	5.1	<0.110	9.300
New Construction	49	0.016	6.5	<0.005	3.200
Other/not specified	14	0.017	4.0	<0.006	0.250
Demolition	4	0.008	22.3	<0.001	0.260

For this study, work tasks were categorized into nine groups based on material and tool used (see Table 5). The most common task measured was cutting and sawing (33%). Demolition had the highest average RCS GM concentration; however, only two samples were from that task. Grinding tasks had the greatest variability in RCS exposure (GSD = 12.4).

Table 5: RCS exposure concentrations by broad work task type (by descending GM)

	N	GM (mg/m^3)	GSD	MIN (mg/m^3)	MAX (mg/m^3)
Demolition	2	0.612	1.3	0.500	0.750
Breaking	19	0.105	10.5	<0.001	9.300
Grinding	14	0.073	12.4	<0.009	3.200
Cutting/Sawing	33	0.055	4.7	<0.010	1.700
Drilling	7	0.026	2.3	<0.021	0.100
Cleaning	7	0.015	3.2	<0.007	0.170
Mixing/Pouring	16	0.010	2.9	<0.007	0.310
Moving and/or Crushing Rocks and/or Earth	21	0.009	4.1	<0.005	0.250
Other	2	0.007	1.0	<0.013	<0.013

The most common material being handled during our exposure monitoring was concrete, though in various forms. Concrete handling gave an average RCS exposure of 0.005 mg/m^3 while the least commonly used material was cement, which also had the highest average RCS concentration (GM 0.083 mg/m^3).

Table 6: RCS exposure concentrations by Material Type (by descending GM)

Material	N	GM (mg/m^3)	GSD	MIN (mg/m^3)	MAX (mg/m^3)
Cement	4	0.083	2.1	0.035	0.160
Concrete	74	0.052	8.8	<0.001	9.300
Asphalt	33	0.039	2.3	<0.041	0.100
Stone	7	0.032	5.5	<0.012	0.250
Gypsum and jointing material	14	0.023	1.7	<0.021	0.056
Other	6	0.008	1.2	<0.013	0.010
Various Material containing sand	13	0.005	2.0	<0.005	<0.028

The work environment also impacts RCS concentration. Fifty-four percent of the samples in this study were collected outdoors, which had the widest range of measurements (GSD 8.2). Samples taken in restricted spaces (N=4) had the highest average concentration compared to other work environment (GM 1.132 mg/m^3).

Table 7: RCS Exposure Levels by Work Environment (by descending GM)

	N	GM (mg/m^3)	GSD	MIN (mg/m^3)	MAX (mg/m^3)
Restricted Space	4	1.132	2.0	0.600	2.900
Indoors	51	0.039	4.2	<0.007	0.750
Outdoors	66	0.024	8.2	<0.001	9.300

Another important factor that influences RCS concentration is the use of controls such as vacuums or wetting. Of all the samples, 77% did not use control strategies (n= 93). The highest average RCS concentration were found in samples that used local exhaust ventilation that is not attached to the work tool (GM 0.230 mg/m^3), although there were only 2 such samples. The lowest average concentration was in samples that used exhaust ventilation and wetting (GM 0.18 mg/m^3, n = 4).

Table 8: RCS Exposure Levels by Control Type (by descending GM)

	N	GM	GSD	MIN	MAX
Local Exhaust Ventilation	2	0.230	1.5	0.17	0.310
Exhaust Ventilation on Tool	21	0.038	5.2	<0.007	1.100
Uncontrolled	93	0.032	7.6	<0.001	9.300
Water Spray on Tool	1	<0.041	NA	<0.041	<0.041
EV+ Wetting	4	0.018	6.0	<0.012	0.250

3.1.1 Provincial Differences

We compared our findings in Manitoba with recent results from similar exposure surveillance done in the provinces of Alberta and British Columbia (Gorman Ng and Davies, 2018).

Manitoba (Table 2) had the lowest overall average concentration for RCS (GM 0.33 mg/m3). The sample variability for RCS (GSD 7.0) was lower than Alberta (GSD 9.3) but higher than BC (GSD 4.3). However, Manitoba had the highest measured RCS concentration across all three provinces (9.3 mg/m^3).

We further investigated these regional differences in a multiple linear regression model, which also allowed us to examine other determinants of exposure to RCS simultaneously (Table 8). The final model created accounted for 37% of the variability in RCS exposure. In this model, Manitoba had one-half the RCS concentration compared to Alberta, and one-third of BC exposure levels, when all other factors such as task, material, environment, controls were controlled for.

Table 9: Multiple Linear Regression Model (adjusted R^2 = 0.37, Total N= 373)

	N	Beta	AntilogBeta (partial F-Test)	P value
Province			(< 0.001)	
Alberta	129	REF	REF	REF
British Columbia	123	0.4097	1.50637	0.124314
Manitoba	121	-0.6502	0.52194	0.004379
Task			(< 0.001)	
Other	35	REF	REF	REF
Breaking	44	2.6308	13.88487	<0.001
Cleaning	22	1.1953	3.30455	1.08E-02
Cutting/Sawing	67	2.511	12.31724	<0.001
Demolition	10	1.9312	6.89778	<0.001
Drilling	26	1.2928	3.64297	0.004599
Grinding	53	2.2632	9.61380	<0.001
Mixing and Pouring	37	0.1815	1.19901	6.74E-01
Moving and/or Crushing Rocks and/or	65	0.7302	2.07550	0.087413
Spraying	7	0.9254	2.52288	0.18337
Environment			(< 0.001)	
In	194	REF	REF	REF
Out	158	-0.4839	0.61637	0.022691
Restricted Space	21	1.519	4.56766	<0.001
Material			(< 0.001)	
Not Specified	10	REF	REF	REF
Other (brick, ceramic, mortar and	29	0.4971	1.64395	4.70E-01
Asphalt	14	-0.6081	0.54438	0.419584
Concrete	191	0.2185	1.24421	0.72208
Cement	38	0.6918	1.99731	0.265915
Sand	46	0.8357	2.30643	0.177637
Gypsum and joint material	31	-1.5222	0.21823	0.022463
Stone and Granite	14	0.9259	2.52414	0.183725
Sampling Duration Category			(< 0.001)	
0.32 -103 minutes	280	REF	REF	REF
104- 240 minutes	32	-1.1835	0.30621	<0.001
241- 390 minutes	44	-0.8974	0.40763	<0.001
391- 12480 minutes	17	-0.9769	0.37648	0.031432
CSP Control			(< 0.001)	
Uncontrolled	240	REF	REF	REF
LEV	37	-0.839	0.43214	0.008548
EV on tool	62	-0.8584	0.42384	0.002271
Water spray on tool	12	-0.6287	0.53328	0.248427
EV+ Wetting	22	-1.0977	0.33363	0.015507

3.2 Risk Assessment Tool Updates

3.2.1 Improvement of the current SCT prediction model

As described in the methods, the updated prediction calculation and associated uncertainty are now based on matrix calculation and multi-model inference. Appendix 3 describes the calculations in detail. In brief, one complication is that the model is fit to 20 datasets. This is due to the format of the historical database, where several results are not individual datapoints but summary results. As described in Lavoué *et al*. (2007) and Sauvé *et al*. (2012 & 2013) these summary results were used to estimate individual measurements using Monte Carlo simulation. For example, for a study reporting exposure in the form of GM and GSD for 15 measurements, we would simulate 15 random values from a distribution with the same GM and GSD. The simulated values are thereafter mixed with the individual values for the modelling. In order to take into account simulation variability, the simulation was performed 20 times, hence the model has to be fit to 20 datasets. The predictions for a particular scenario therefore have to be averaged across the 20 model fits, and variability across the 20 models needs to be taken into account in the final uncertainty estimates.

3.2.2 Improvement of the risk metrics calculated from the SCT model

Figure 3 below shows how an example use of the risk assessment framework from British Columbia was implemented, where the main risk metric is the geometric mean (GM) of the distribution of workers' exposures. Risk is considered controlled if the GM is below the OEL (in this case, we are using an OEL of 0.1 mg/m^3). The figure is a screenshot from the prototype web application developed during this project.

The data analysed is a simulated dataset of 9 silica measurements: 0.070 / 0.082 / 0.080 / 0.025 / 0.081 / 0.025 / 0.050 / 0.055 / 0.006 mg/m^3.

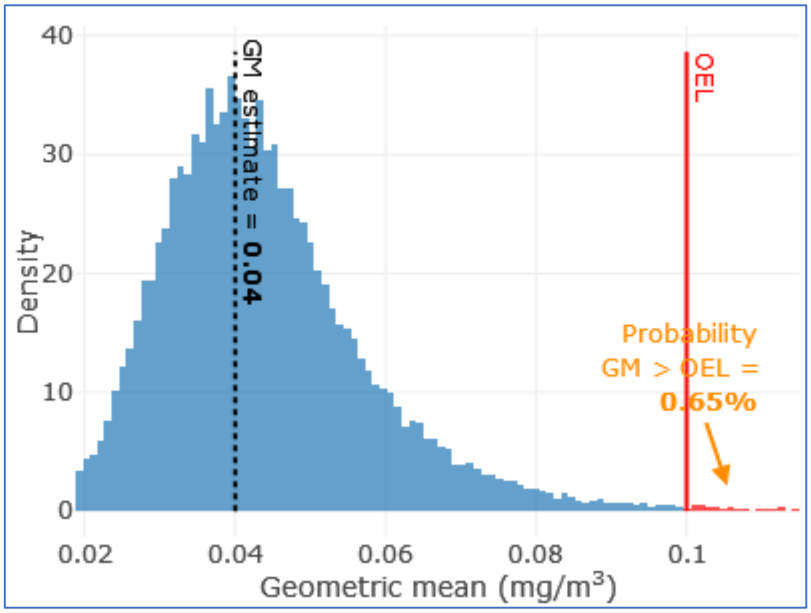

Figure 3: plot output of the new silica prototype for a simulated dataset (analysis according to the British Columbia risk analysis framework).

The point estimate of the geometric mean is 0.04 mg/m^3, with a 95% upper confidence limit of 0.07 mg/m^3. The plot shows the uncertainty distribution of the geometric mean, i.e., a histogram of 25,000 possible values for the GM generated by the Bayesian model. The plot also illustrates the probability that the true GM is above the OEL, i.e., the proportion of the histogram above the OEL: 0.65%. In this case we are very confident that the GM of the exposure distribution is well below the OEL of 0.1 mg/m^3.

Figure 4 below shows the results of the analysis using the second available risk analysis framework, advocated by the AIHA, where the main risk metric is the 95th percentile (P95).

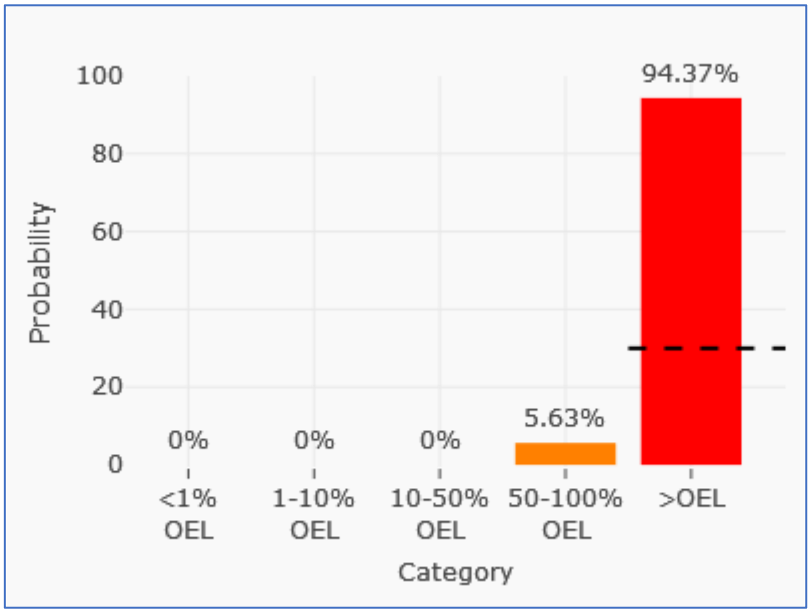

Figure 4: plot output of the new silica prototype for a simulated dataset (analysis according to the AIHA analysis framework).

The point estimate of the 95th percentile of the exposure distribution is 0.18 mg/m^3, above the OEL, and the corresponding 95% upper confidence limit is 0.64 mg/m^3. The graph in Figure 4, which splits the uncertainty distribution of the 95th percentile into 5 risk management categories, shows, unsurprisingly, a very high probability (94.4%) of the situation corresponding to a true 95th percentile above the OEL (of 0.1 mg/m^3). In this case we are therefore confident that the situation represents unacceptable exposure in this framework.

3.2.3 Bayesian model for combining the model prediction with a "new" silica dataset

As mentioned in the Methods section, we adapted one of the models described in the WEBEXPO scientific reports which serves as the backbone of the Expostats data interpretation toolset (Lavoué *et al.*, 2018). Appendix 4 describes the mathematics of the new model. In essence, the uninformative prior on the geometric mean in the Expostats model (InformedVar in Appendix 4) was modified for the purpose of this project. Instead of being uniform (in the log scale) over a wide range, the new prior is in the shape of a normal distribution defined by a mean and standard deviation, which matches exactly the improved output of the silica prediction model (see 3.2.1 and Appendix 3). A technical web application was first built that ran this model and the normal Expostats model concurrently, to help the research team explore the influence of the new information on the final exposure estimates. This test app is available at https://lavoue.shinyapps.io/SilicaBC/.

As an illustration, Figure 5 below shows an example comparing the uncertainty distribution around the geometric mean (in a similar fashion as Figure 1) as predicted by the prior alone (the silica prediction model) in the top part, a simulated "new" dataset (n=9 samples) alone (obtained through the Expostats model), and the posterior distribution obtained using the new Bayesian model, which *combines* information from the prediction model and the new data. In this example the prior information suggest a very low GM compared to the OEL (using an example OEL of 0.15). In the middle part of the graph, the 9 measurements suggest a somewhat higher GM value, although still low compared to the OEL. The final combined uncertainty distribution illustrates the Bayesian process of updating the prior information with the new data, the final (posterior) information clearly appearing an "average" of the 2 other distributions. As both the "prior" and "data" uncertainty distributions for the GM reflect similar uncertainty (both are fairly flat and spread across a wide range), the averaging seems quite balanced (i.e., both distributions have a similar influence on the final estimate).

Figure 6 shows the same process but this time the sample size for the new data was 50. The greater sample size is reflected in the "data" distribution, this time reflecting much lower uncertainty about GM than the prior (much narrower histogram). As a consequence, the posterior distribution is almost equal to the "data" distribution. This is an illustration of data overwhelming the prior.

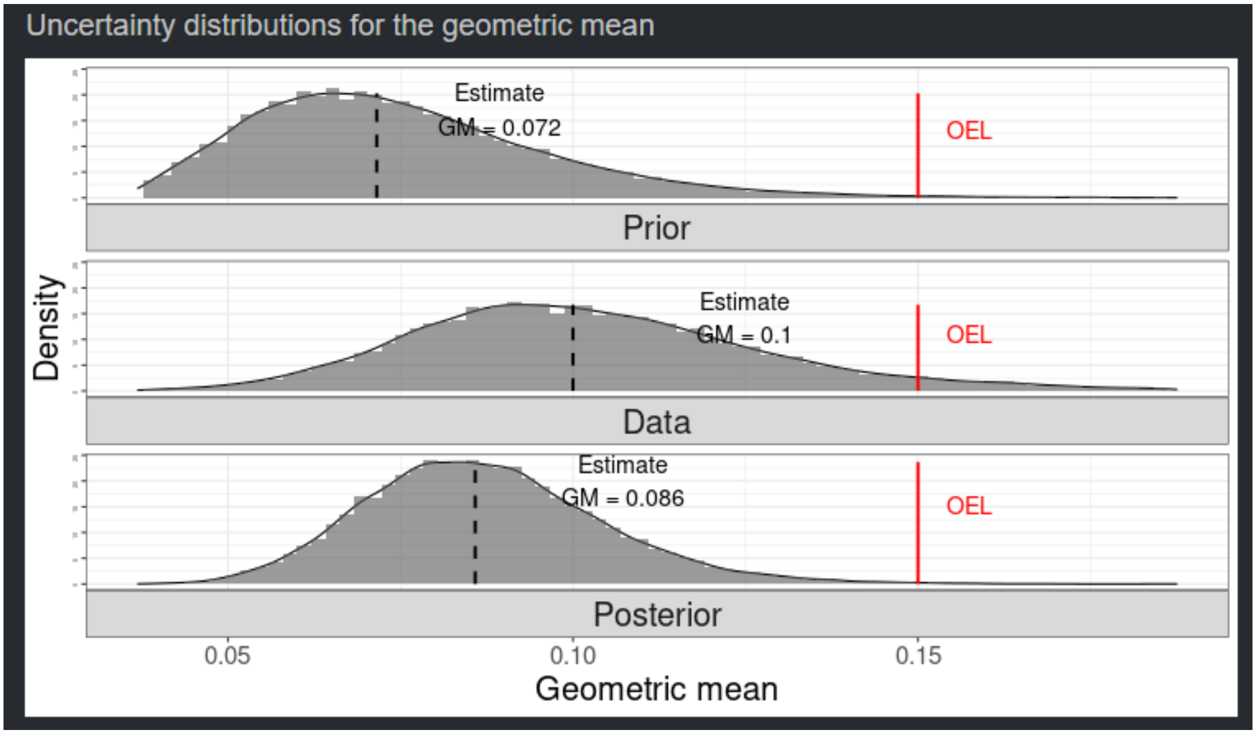

Figure 5: Example of updating an informed prior distribution with new data (n=9).

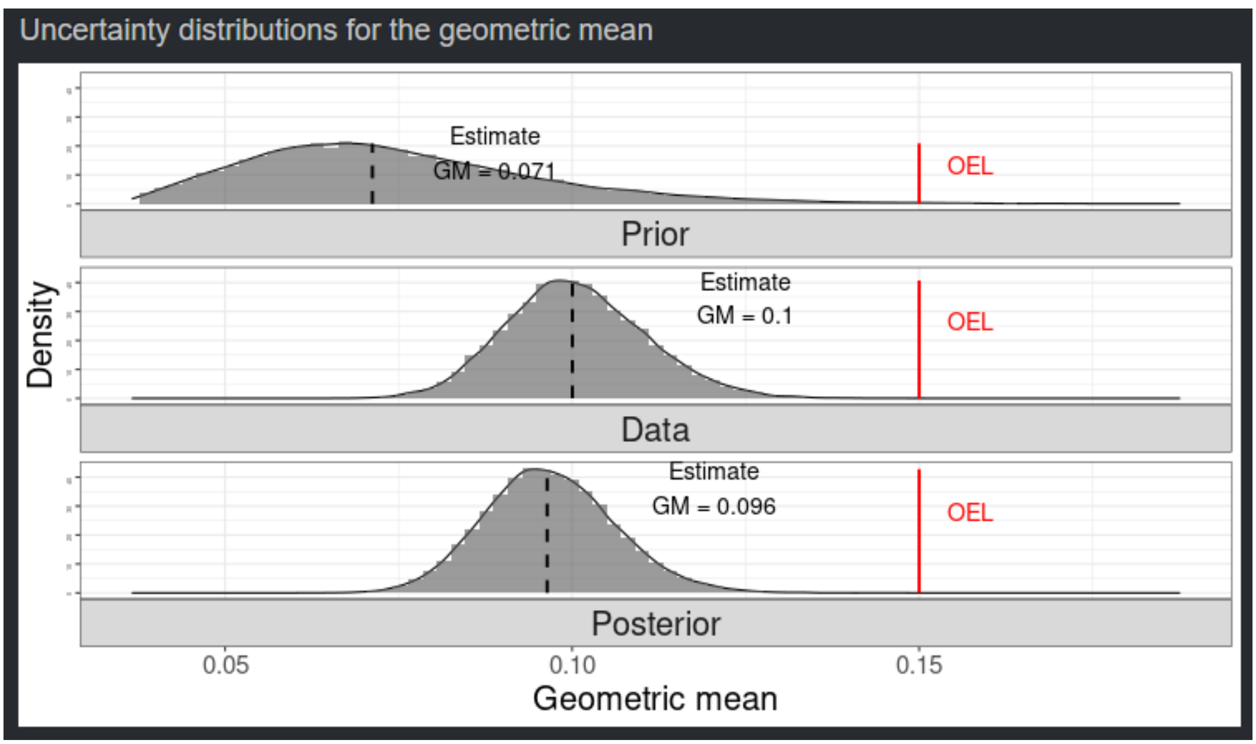

Figure 6: Example of updating an informed prior distribution with new data (n=50).

3.2.4 Developing improved user interface

The silica prediction model and the Bayesian models described above were integrated in a javascript web prototype for a risk assessment tool providing results according to both the British Columbia and AIHA risk decision framework. One aim was to allow users to see the results from the silica model prediction alone, from a "new" datas*et al*one, and from their combination using the newly available Bayesian model.

Figures 7 and 8 below showcase this prototype, now fully functional and available from https://silica.expostats.ca/. Figure 7 shows the main entry point, where the first step consists in selecting a prediction scenario using the exposure determinants included in the Silica model. After selecting the scenario and the OEL, the user is shown the results of both risk decision frameworks (BC and AIHA) according to the *historical* database. In step 2, the user can enter their own measurement data, and the Expostats Bayesian model is run to show the risk analysis based only on *just these* data. In step 3, the user is shown risk analysis results from the *combination* of both sources of information. Finally screen 4 (shown in figure 8) is a recap of the three steps, when users can see the effect of updating the silica prediction model estimate with the new data as shown, e.g., in Figure 3 and 4.

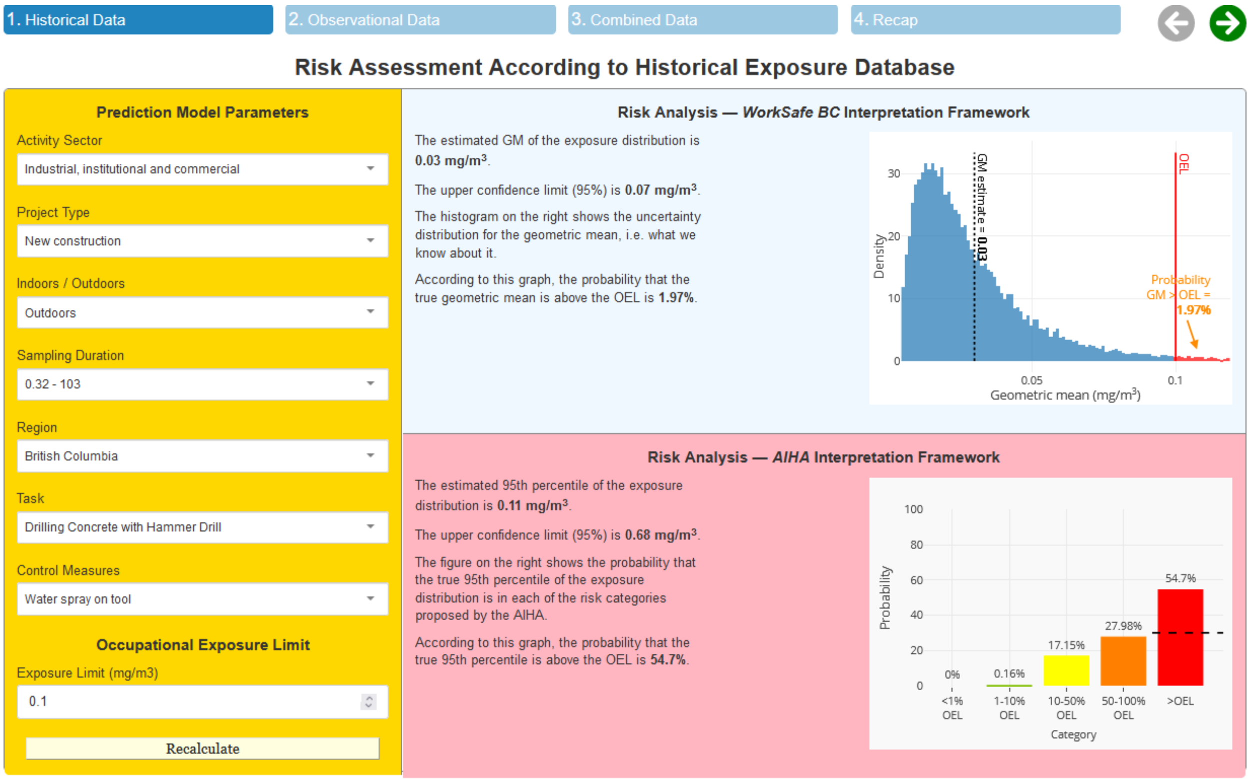

Figure 7: Entry point of the Silica risk prediction prototype

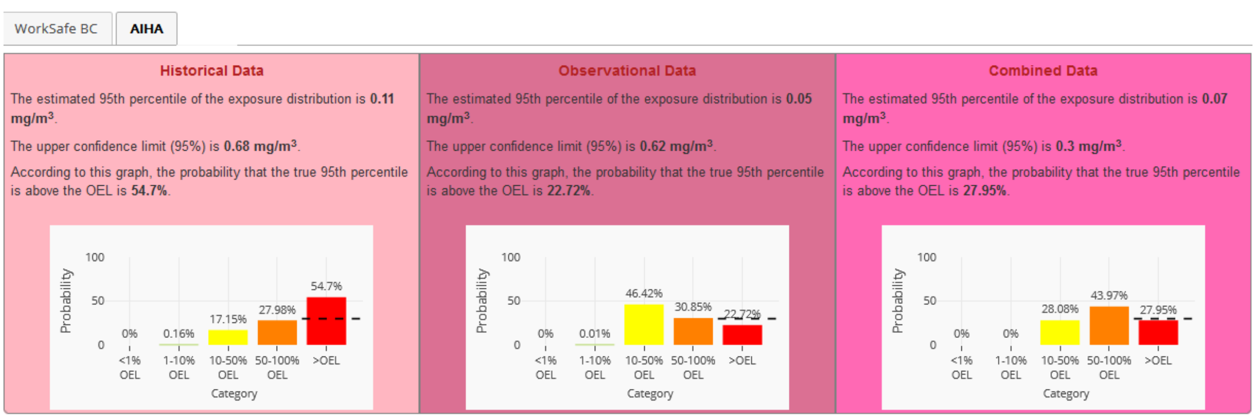

Figure 8: Recap screen of the Silica risk prediction prototype for the AIHA risk decision framework

4 Discussion

4.1 Respirable Crystalline Silica Exposure in Manitoba Construction Industry

This study found levels of RCS in MB construction industry that were similar levels found in contemporary studies in Alberta and British Columbia. The mean level found in MB (geometric mean = 0.033 mg/m^3) was the lowest of the three provinces but the mean still exceeded the Manitoba OEL (and the health-based ACGIH TLV-TWA®) of 0.025 mg/m^3.

The variability in the 121 personal exposure measurements (GSD = 7.0) indicates that the measurement observations are strongly skewed, A few measurements had very high concentrations. The maximum value observed (9.3 mg/m^3, while jackhammering dry concrete without dust controls) was the highest of any measurement in recent sampling campaigns by the authors in Alberta, BC or Manitoba. Such high levels (approximately 300 times above the ACGIH TLV-TWA®) are concerning. There have been recent outbreaks of silicosis and other respiratory symptoms associated with relatively short exposures to similarly high levels of RCS, and acute silicosis (also known as silicoproteinosis) can occur within 1-10 years of higher exposures in the range of 1-10mg/m^3/year (Tustin, 2022; Barnes *et al*, 2019; Rose *et al.*, 2019).

Nine of the 14 companies sampled had (geometric) mean exposure levels over the TLV-TWA® of 0.025 mg/m^3. Note that this should only be used as a relative indicator of the degree of the observed exposure levels and not a statistically rigorous compliance test; nevertheless, it does indicate that over-exposure to RCS is likely a wide-spread problem in construction in Manitoba.

Our exposure measurements were short-term, task-based samples of average duration of 55 minutes. If employee exposures during a typical shift are lower (for example because employees undertake a variety of tasks, some with lower exposure levels) then the actual 8-hour average exposure may be lower and thus our estimates may be considered to be over-estimates and therefore overly protective. However, over protection should not be assumed. Many construction projects are completed by specialty subcontractors who may undertake the same tasks repeatedly for long periods throughout the day. Furthermore, there are many construction activities that generate RCS so background exposure levels are often greater than zero.

Our measurements were all obtained on active worksites during normal construction activities, and therefore should be broadly representative of conditions expected at any Manitoba workplace. It is possible that employers who volunteered to participate may in fact be considered more likely to be better informed about OHS hazards and silica, and therefore have "cleaner" sites than the average. This would then mean that our observed values might underestimate true RCS exposures in Manitoba construction sector.

4.2 Determinants of exposure

We used linear regression analysis to examine the factors influencing RCS exposure levels on worksites. Variable examined were province, work task, environmental setting (indoor/outdoor, etc.), construction material, duration of measurement, and type of dust control. The modeling we performed explained approximately 40% of the observed variability. The estimates in Table 9 show the relative difference from a base model; for example, compared to Alberta (the reference province) and holding all other variables constant, BC RCS measurements were 50% greater while MB measurements were on average half of Alberta levels.

4.2.1 Regional differences

Some of the differences observed between provinces may be due to systematic differences in the sampling conditions between provinces (for example, more MB measurements were made outdoors, or in MB we measured task types producing less RCS dust). Many of these factors were included in the regression model and thus "accounted for". Because the measurement of these factors was imperfect, some of the observed variability likely still comes from these factors, however some will also be coming from factors we did not measure and can be considered to likely vary on a regional basis.

One source of regional difference may be differences in the basic silica-containing materials. For example, the material involved during various construction tasks may be different (e.g., gypsum vs sand lime) or the formulation of material may vary (percent of sand in a particular concrete mix) and/or the geographic source of material due to regional geological variation.

The formulation of material may contribute to differences in RCS released. For example, most concrete is composed of sand, aggregate rocks and paste (cement mixed with water) where sand and aggregate rocks usually contain high silica content. An industrial project using concrete that is 30% sand, 40% aggregate and 30% paste would release different amount of silica if the formulation and percentage of sand and aggregate changes. Higher sand and aggregate percentage would likely lead to higher RCS concentration. Additionally, industrial projects process concrete mixture for longer; thus, pushing the sand and aggregate into lower layers and have more paste on the surface of the material, which contain little to no silica (Flanagan *et al*., 2006).

Using the same concrete example as above, silica content may differ depending on which type of sand and aggregate rocks are used and where they are sourced. We expect that construction companies in Alberta, BC and Manitoba would source their material locally and the local geological environment may produce different types of rocks. For example, rocks in BC are composed of volcanic, intrusive and some metaphoric rocks. Manitoba has mostly sedimentary rocks with some intrusive and little volcanic rocks and Alberta has mostly sedimentary rocks (Government of Canada Atlas of Minerals and Mining, 2020). High percentages of crystalline silica quartz are commonly found in sedimentary rocks (ie. sandstones) and metamorphic rocks, which are more typical of rock types in Alberta and Manitoba (Carmichael, 1989; Atkinson & Atkinson, 1978; US Bureau of Mines, 1992; Heaney & Banfield, 1993; Ross *et al*., 1993; IARC,1997 p65). This is consistent with our study where Alberta had the highest average RCS concentration compared to BC and Manitoba. However, Manitoba had the lowest average RCS concentration which is inconsistent with the expectation that metamorphic rocks have high concentration of RCS.

4.2.2 Other sources of variability

We found exposure level varied by construction project type. Other studies have also found that demolition tasks are associated with high RCS exposure relative to other construction tasks. Demolition often involves various materials using wide range of tools from handheld hammers for tile removal to small bulldozers to push over walls (Lumens & Spee, 2001; Nij *et al.*, 2004; Radnoff *et al.*, 2014).

Task explained most of the variability in RCS concentration. For our study, task type is defined by the combination of task, tools, and material, which are all factors that can influence RCS exposure. Other studies have also reported similar breaking tasks having high RCS concentration (Flanagan *et al.*, 2006; Si *et al.*, 2016). Grinding may be highly variable due to different types of tools and technique used such that the amount of RCS released may be determined by grinder diameter, surface area, wheel type and rotation rate. For example, Flanagan *et al.* (2003) showed that a 4.5-inch grinder released 33% less RCS than a 7-inch grinder while abrasive wheel released 60% less RCS compared to a diamond wheel. The grinding technique also mattered such that side-to-side movement and excessive force used on grinder may release more RCS from the material during mortar removal and renovation (Flynn & Susi, 2003).

Regression modeling (Table 9) shows that engineering controls such as LEV and wetting of surfaces or tools can have a significant impact on exposure levels reducing them on the order of one-half.

4.3 Improving risk assessment tools for RCS

Occupational hygienists widely agree on the importance of quantitative exposure measurement (Kromhout, 2016) for compliance monitoring, but also control selection, intervention evaluation, surveillance of trends in exposure and for education purposes. Yet, even where exposure data is available, data interpretation techniques remain complex, and largely in the domain of a subset of industrial hygienists (Lee *et al*., 2019). The authors previously addressed this in British Columbia by developing the "Silica Control Tool" with the aim of providing timely, quantitative risk assessments for complex work environments, and combining ease-of-use for non-experts, cost-effective data collection, data mobilization, and to drive continuous improvement in terms of controls (Davies and Gorman-Ng, 2020).

The limitations of the online tool presented opportunities for improvement. In particular, the original statistical modeling and estimating tool, while adequate, did not provide a means to incorporate new exposure measurement data into the database in a way that efficiently utilizes both old (available in quantity, less specific) and new data (very few measurements but more specific). In addition, typical occupational hygiene exposure metrics (geometric mean, geometric standard deviation, 95th percentile) are often difficult to interpret even for OH professionals.

This project has significantly improved the core functioning of the tool by (i) incorporating Manitoba-specific data to the underlying database used in modelling and predicting RCS exposures; (ii) re-coding the core statistical processes to make them more efficient and robust; (iii) enable the end-user to enter their own exposure data and have it modelled "in context" of the wealth of historical data and (iv) create a new results dashboard with multiple risk-assessment formats.

The ability to combine modeling predictions with exposure data has been identified as a limitation of current methods and tools and this study has contributed significantly to overcoming this problem in a practical manner (Ramachandran, 2019). These technical improvements are ready to be migrated into the BCCSA silica control tool in a future project. The improvements also offer tremendous potential to be generalized as an end-user tool for any occupational exposure database, for example the Canadian Workplace Exposure Database, or CWED (https://cwed.spph.ubc.ca/).

4.4 Review of work completed

The project met its overall aim of improving understanding of RCS hazard in Manitoba construction industry by:

1. Undertaking a large exposure monitoring campaign over two construction seasons. This comprised (see Section 3.1):
   - Identifying appropriate "common silica-producing processes" for Manitoba industry (for planning and prioritizing sampling strategy)
   - Recruiting fourteen MB construction companies
   - Obtain, analyze and report results from 121 personal respirable crystalline silica (RCS) exposure samples
2. Improving RCS exposure data holdings vis-à-vis regional, contemporary and new exposure scenarios. This comprised (see Section 3.1.1):
   - RCS exposure data base available from authors on request
   - Incorporated MB RCS exposure data into BCCSA SCT database
3. Characterizing RCS exposure situation in MB construction, and comparing to other Canadian regions. This comprised (see section 3.1.1):
   - Data cleaning and coding
   - Undertook statistical analysis to examine the determinants of RCS exposure levels including region
4. Improve modeling and estimation methodology. This comprised (see section 3.2.1 – 3.2.4):
   - Improvement of the current SCT prediction model
   - Improvement of the risk metrics calculated from the SCT model
   - Bayesian model for combining the model prediction with a "new" silica dataset
   - Developing improved user interface

4.5 Recommendations

The findings of the study suggest that the levels of RCS to which Manitoba construction workers may be exposed are consistent with levels observed in other region of Canada, and other countries. The findings, taken in context of other work and general OHS principles led to a series of recommendations:

4.5.1 RCS safety: recommendation to regulator, employers, safety associations

Referring to OHS hierarchy of controls (Figure 9), there are clearly opportunities for controlling exposure from RCS at all levels. Our study did not record or report on PPE use in study participants but the limitations of PPE (in this case, respiratory protection) is well known. When working with RCS if PPE is to be used it should be as part of a well-managed respirator program that ensures correct type of respirator is used, it is well-fitted, maintained and worn when needed.

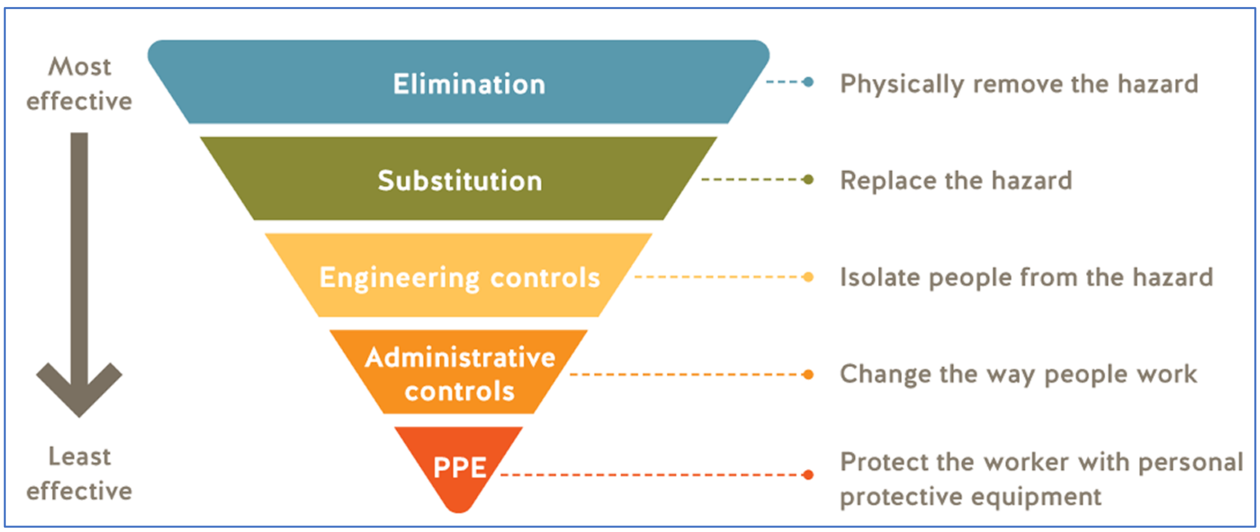

Figure 9: The OHS Hierarchy of Controls (courtesy WorkSafeBC Website)

With respect to eliminating the RCS exposure hazard, while it is challenging to eliminate the materials containing crystalline silica there may be opportunities to eliminate the tasks that produce high levels of RCS exposure. In our study that might mean trying to reduce demolition tasks, breaking, and grinding. This might be achieved through changes in how work is designed – like using better concrete forms to reduce the amount of grinding required.

Our findings showed the benefit of tool-based (engineering) controls (wetting and exhaust ventilation) when used in tandem; however, use of controls mounted on tools did not result in greatly reduced exposure levels, when used individually. This may point to the fact that these engineering solutions also require proper specification, maintenance and use to be effective. Use of general local exhaust ventilation was associated with an increase in exposure level in our study but there were only two measurements on which to base this finding and so they could be non-representative. Further, LEV may have been used in this instance because exposure levels were expected to be high.

Vis-à-vis administrative controls, the findings of this study are useful to inform training materials for employees, OHS specialists and management around RCS hazards and controls. Further, quantitative exposure estimates using data such as that available from this study allows for planning of job-rotation that can limit average exposures over time.

4.5.2 Knowledge Translation: recommendations for researchers, regulators

The findings of this study can be used to inform educational materials with respect to: (i) silica health hazard; (ii) hazardous exposures to RCS; (iii) controlling RCS exposure. Target audiences include employees, employers, apprenticeship trainers, health and safety regulatory officers, claims adjudicators and others in compensation and prevention.

The findings of this study should be presented in two scientific papers: (i) on the exposure assessment of RCS in Manitoba (a draft manuscript has already been prepared by the authors) and (ii) a scientific paper describing the process of developing the enhanced statistical "engine" for the Silica control Tool, and the new end-user "dashboard".

The researchers should report the findings of the study to the Manitoba construction Industry and other stakeholders in 2022 at an appropriate construction meeting (for example the CSAM safety conference).

4.5.3 Silica Control Tool upgrades

Upgrades to the computer programming of the BCCSA Silica Control Tool (https://lavoue.shinyapps.io/SilicaBC/) as well as the new end-user interfaces for risk assessment (https://silica.expostats.ca/) should be integrated into the Tool. The researchers will seek additional grant funding for this step.

References

Atkinson, F. & Atkinson, R. (1978) The Observer's Book of Rocks and Minerals, Claremont, GA, United States, Claremont Books

Beaudry, C., Lavoué, J., Sauvé, J.-F., Bégin, D., Senhaji Rhazi, M., Perrault, G., Dion, C., & Gérin, M. (2013). Occupational exposure to silica in construction workers: A literature-based exposure database. *Journal of Occupational and Environmental Hygiene*, *10*(2), 71–77. https://doi.org/10.1080/15459624.2012.747399

Banerjee, S., Ramachandran, G., Vadali, M., & Sahmel, J. (2014). Bayesian Hierarchical Framework for Occupational Hygiene Decision Making. *Annals of Occupational Hygiene*, *58*(9), 1079-1093. https://doi.org/10.1093/annhyg/meu060

Barnes H, Goh NSL, Leong TL, Hoy R. Silica-associated lung disease: An old-world exposure in modern industries. Respirology. 2019 Dec;24(12):1165-1175. doi: 10.1111/resp.13695. Epub 2019 Sep 13. PMID: 31517432.

Building Force Canada. (2021). Construction and Maintenance Looking Forward Manitoba employment slows from peak demands Highlights 2021-2030. Availible from https://www.buildforce.ca/system/files/forecast_summary_reports/2021%20M%20Constr%20Maint%20Looking%20Forward.pdf. Accessed 23 December 2021

Carmchael, R.S. (1989) Practical Handbook of Physical Properties of Rocks and Minerals, Boca Raton, FL, CRC Press

CAREX Canada (2021). Silica (Crystalline) Occupational Exposures. Downloaded from https://www.carexcanada.ca/profile/silica_crystalline-occupational-exposures/; Dec 2021

Cherrie JW, C. Sewell, P. Ritchie , C. McIntosh, J. Tickner & D. Llewellyn (2001) Retrospective Collection of Exposure Data from Industry: Results from a Feasibility Study in the United Kingdom, Applied Occupational and Environmental Hygiene, 16:2,144-148,

Croteau, G. A., Flanagan, M. E., Camp, J., & Seixas, N. (2004). The Efficacy of Local Exhaust Ventilation for Controlling Dust Exposures During Concrete Surface Grinding. *The Annals of Occupational Hygiene*. https://doi.org/10.1093/annhyg/meh050

Davies, H. W., & Gorman-Ng, M. (2020). Development of a Web-Based Tool for Risk Assessment and Exposure Control Planning of Silica-Producing Tasks in the Construction Sector. *Frontiers in Public Health*, *8*, 371. https://doi.org/10.3389/fpubh.2020.00371

Flanagan, M. E., Seixas, N., Becker, P., Takacs, B., & Camp, J. (2006). Silica Exposure on Construction Sites: Results of an Exposure Monitoring Data Compilation Project. *Journal of Occupational and Environmental Hygiene*, *3*(3), 144–152. https://doi.org/10.1080/15459620500526552

Flanagan, M. E., Seixas, N., Majar, M., Camp, J., & Morgan, M. (2003). Silica Dust Exposures During Selected Construction Activities. *AIHA Journal*, *64*(3), 319–328. https://doi.org/10.1080/15428110308984823

Gorman Ng, M. & Davies, H. W. (2018). Evaluating a Respirable Crystalline Silica Risk Assessment Model for the Construction Industry in Alberta; Report to Government of Alberta OHS Futures – Research Funding Program

Hall, A, C. Peters, P Demers, HW DAvies (2014) Exposed! Or not? The diminishing record of workplace exposure in Canada; Canadian journal of public health. Revue canadienne de santé publique 105(3):e214-e217

Heaney, P.J. & Banfield, lA (1993) Structure and chemistry of silica, metal oxides, and phosphates. In: Guthrie, G.D. & Mossman, B.T., eds, Reviews in Mineralogy, Vol. 28, Health Effects of Mineral Dusts, Chelsea, MI, Book Crafters, 185-233

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

IARC. (1996). Silica, Some Silicates, Coal Dust and Para-Aramid Fibrils. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 68.* Lyon (Fr): International Agency for Research on Cancer. ISBN 92 832 1268 1

IARC. (2012). Silica Dust, Crystalline, In the Form of Quartz or Cristobalite. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 100C*. Lyon (Fr): International Agency for Research on Cancer. Available at https://www.ncbi.nlm.nih.gov/books/NBK304370/. Accessed 12 November 2021

Jahn, S. D., Bullock, C., & Ignacio, J. S. (2015). *A Strategy for Assessing and Managing Occupational Exposures—4th edition*. AIHA Press.

Kromhout, H. (2016). Hygiene Without Numbers. *Annals of Occupational Hygiene*, *60*(4), 403–404. https://doi.org/10.1093/annhyg/mev096

Lavoué, J., Joseph, L., Knott, P., Davies, H., Labrèche, F., Clerc, F., Mater, G., & Kirkham, T. (2018). Expostats : A Bayesian Toolkit to Aid the Interpretation of Occupational Exposure Measurements. *Annals of Work Exposures and Health*. https://doi.org/10.1093/annweh/wxy100

Lavoué, J., Bégin, D., Beaudry, C., & Gérin, M. (2007). Monte Carlo simulation toreconstruct formaldehyde exposure levels from summary parameters reported in the literature. *The Annals of occupational hygiene*, *51*(2), 161-172. https://doi.org/10.1093/annhyg/mel068

Lee EG, Lamb J, Savic N, Basinas I, Gasic B, Jung C, et al. Evaluation of exposure assessment tools under REACH: part I-tier 1 tools. Ann Work Expo Health. (2019) 63:218–29. doi: 10.1093/annweh/wxy091

NIOSH (2002). Health effects of occupational exposure to respirable crystalline silica. U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health. https://doi.org/10.26616/NIOSHPUB2002129

OCRC (2019) Burden of Occupational Cancer in Canada; Report by the Occupational Cancer Research Center, Toronto, Ontario; available at https://www.occupationalcancer.ca/2019/national-burden-report/

Quick, H., Huynh, T., & 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. https://doi.org/10.1093/annweh/wxw001

Radnoff, D., Todor, M. S., & Beach, J. (2014). Occupational Exposure to Crystalline Silica at Alberta Work Sites. *Journal of Occupational and Environmental Hygiene*, *11*(9), 557–570. https://doi.org/10.1080/15459624.2014.887205

Radnoff, D., Todor, M. S., & Beach, J. (2015). Exposure to Crystalline Silica at Alberta Work Sites: Review of Controls. *Journal of Occupational and Environmental Hygiene*, *12*(6), 393–403. https://doi.org/10.1080/15459624.2015.1009987

Ramachandran, G, (2019) Progress in Bayesian Statistical Applications in Exposure Assessment; Annals of Work Exposures and Health, 2019, Vol. 63, No. 3, 259–262

Rappaport, S. M., Goldberg, M., Susi, P., & Herrick, R. F. (2002). Excessive Exposure to Silica in the US Construction Industry. *The Annals of Occupational Hygiene*. https://doi.org/10.1093/annhyg/meg025

Rose, C, et al. (2019) Severe Silicosis in Engineered Stone Fabrication Workers — California, Colorado, Texas, and Washington, 2017–2019; MMWR; Vol 68 / No. 38

Ross, M., Nolan, R.P., Langer, AM. & Cooper, W.C. (1993) Health effects of mineral dusts other than asbestos. In: Guthrie, G.D. & Mossman, B.T., eds, Reviews in Mineralogy, VoL. 28, Health Effeets of Mineral Dusts, Mineralogical Society of America, Chelsa MI, Book Crafters, pp. 361-407

Sauvé, J.-F., Beaudry, C., Bégin, D., Dion, C., Gérin, M., & Lavoué, J. (2012). Silica Exposure During Construction Activities: Statistical Modeling of Task-Based Measurements from the Literature. *The Annals of Occupational Hygiene*. https://doi.org/10.1093/annhyg/mes089

Sauvé, J.-F., Beaudry, C., Bégin, D., Dion, C., Gérin, M., & Lavoué, J. (2012). Statistical modeling of crystalline silica exposure by trade in the construction industry usinga database compiled from the literature. *Journal of Environmental Monitoring : JEM*, *14*(9), 2512-2520. https://doi.org/10.1039/c2em30443k

Statistics Canada. Table 14-10-0092-01 Employment by industry, annual, provinces and economic regions, inactive (x 1,000) **DOI:** https://doi.org/10.25318/1410009201-eng

Stacey P, Thorpe A. (2010) Testing of high flow rate respirable samplers to assess the technical feasibility of measuring 0.05 mg.m^{-3} respirable crystalline silica: (Research Report RR825). Sudbury, Suffolk, UK; HSE Books. Available at https://www.hse.gov.uk/research/rrpdf/rr825.pdf. Accessed 18 December 2021

United States Bureau of Mines (1992) Crystallne Silica Primer, Washington DC, United States Department of the Interior

WorkSafeBC (2021) OHS Regulations Part 5 Guidelines. Available from https://www.worksafebc.com/en/law-policy/occupational-health-safety/searchable-ohs-regulation/ohs-guidelines/guidelines-part-05; Downloaded 12/1/21

Zhang, Y. F., Banerjee, S., Yang, R., Lungu, C., & Ramachandran, G. (2009). Bayesian modeling of exposure and airflow using two-zone models. *The Annals of occupational hygiene*, *53*(4), 409-424. https://doi.org/10.1093/annhyg/mep017

Appendix 1 Common Silica Processes

CSP according to the data dictionary that combines task, tool and material

Code	Description
1	Not Specified/Other
2	Cutting Asphalt with walk-behind saw
3	Milling Asphalt with milling machine
4	Cutting concrete masonry unit with water-fed table saw
5	Cutting concrete masonry unit with portable saw
6	Cutting concrete with saw
7	Coring concrete with coring machine
8	Drilling concrete with electric hammer drill
9	Grinding concrete with surface, angle or flat grinder
10	Grinding concrete with counterbalanced ceiling grinder
11	Grinding, Preparing and finishing concrete - other
12	Scarifying or bush hammering (concrete)
13	Demolition (any material)
14	Sweeping (any construction area)
15	Shot-creting
16	Cutting tiles with powered tile saw
17	Manual moving of small rocks, soil, etc
18	Mechanized moving of rocks, soil, etc.
19	Crushing and processing rock/sand/earth
20	Cutting marble/granite
21	Cutting drywall
22	Sanding drywall
23	Mixing and pouring cementicious material
24	Tuckpoint grinding
25	Spraying
26	Other masonry-related tasks
27	Other roadwork
28	Breaking concrete with jackhammer
29	Breaking - other
30	Other cutting
31	Abrasive blasting
32	Tunnel boring
33	Installation of acoustic ceiling tiles
34	Other surface grinding
35	Installation of concrete formwork

36	Other cleaning
37	Other drilling
38	Foundation work
39	Electrical maintenance work
40	Excavation work
41	Cutting fiber cement board with portable saw
42	Loading Concrete mixer truck
43	Crushing and processing Concrete
44	Cutting concrete with walk-behind saw
45	Chipping Concrete
46	Mixing Gypcrete
47	Walk-behind concrete grinding
48	Power Sweeping Concrete
49	Concrete Breaking with Excavator
50	Drilling Rock
51	Asphalt breaking with Excavator
52	Cement plant helper
53	Cement Plant Operator
54	Cement Loader operator
55	Cement truck driver
56	Cement Plant Mechanic
57	Cement Plant Lead Hand
58	Asphalt Plant Operator
59	Asphalt Loader Operator
60	Asphalt Truck loading
61	Asphalt Plant Helper
62	Concrete Breaking with Excavator with Jackhammer Attachment
63	Cutting concrete with electric wire saw
64	Indoor mini batch plant operator
65	Sweeping drywall
66	Shotblasting
67	Hosing during concrete demolition with excavator
68	Pinning Wood to Concrete
69	Screeding Concrete with Gas-powered screed
70	Concrete floor scraping
71	Cutting Artificial Stone Countertops
72	Polishing Artificial Stone and/or Granite Countertops
73	Grinding Artificil Stone Countertops
74	Vacuuming Drywall

75	Mechanical Street Sweeping
76	Cleaning vacuum filters
77	Concrete Drilling (Overhead)
78	Concrete dowel drilling

Appendix 2: Worksite Data Collection Form

Field Form Observation Sheet
Site Location:
Date:
Sample ID:
Common Silica Process (CSP):
Worker Job Title:
Shift Length:
Task Duration: Regular Task Duration: How many times task takes place during a normal shift:
General Observations/Comments:
Engineering Controls in place: YES or NO (circle one) - Sample for the CSP is "controlled" or "uncontrolled" (circle one)
Engineering Controls Description and Observations (i.e., water used, or general LEV, or LEV on tool - make/model/description/age):
Sub-Tasks Observed during the CSP:
Materials Involved in CSP:
Materials MSDS Available: YES or NO (circle one); If YES, silica %:
Tools Used (make/model/description):
PPE Worn/Used:
Work environment: Temperature: Precipitation: Pressure: Wind: Indoors / Outdoors (circle one)
Environment/Work Area Description (indoors/outdoors, confined space, enclosed, partial, etc…):
Construction type: New build or Renovation (circle one)
Site Category: Residential Industrial Institutional/Commercial Civil/Roadwork Other:
Outside Temperature: Precipitation: Pressure: Wind:
Photos taken: Yes or No Photo ID (if applicable):
Environmental Conditions/Observations:

Appendix 3: Estimating modelling uncertainty across 20 model fits

Notation:

Y is a n*1 vector of log-transformed exposure levels

β is the p*1 vector of model coefficients, with p the number of parameters in the model

X_h is p*1 the design vector for the prediction scenario of interest

y_hat is the predicted mean for the scenario of interest.

Single model uncertainty[1]

Y_hat is given by $y_{hat} = X_h' \boldsymbol{\beta}$ (1)

The variance of Y_hat (prediction error) is $var_{yhat} = \sigma^2(Y_{hat}) = X_h' \sigma^2(\beta) X_h$ (2) , with $\sigma^2(\beta)$ the variance-covariance matrix of the model

Multimodel uncertainty[2]

For k model fits (in this project k=20), equations (1) and (2) above provide k values of y_hat and var_yhat

The multimodel inference framework provides a way to calculate a prediction averaged over the 20 datasets, and, more importantly an estimate of the variability of the final averaged prediction, according to the following equations

Averaged prediction: $\mathrm{yhat}_{\mathrm{average}} = \frac{\sum_{j=1}^{k} yhat_j}{k}$ (3)

Unconditional variance of the prediction:

[1] Neter, J., Kutner, M., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied Linear Statistical Models: Vol. Fourth edi.* WCB McGraw-Hill.
[2] Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference : A practical information-theoretic approach* (2nd éd.). Springer.

$$\sigma^2(yhat_{average}) = \sum_{j=1}^{k} \frac{\left\{\sigma^2(yhat_j) + \left(yhat_j - yhat_{average}\right)^2\right\}}{k}$$

The standard error on the average prediction is simply the square root of the unconditional variance above.

Appendix 4: Mathematics underlying the Bayesian models

1 Introduction

This document describes the mathematical basis for the two bayesian models used within this project. The `InformedVar` model, used to analyse measurement data with a generic uninformative prior, was previously developped within the Webexpo project. The `InformedMean` model, developped specifically for this project, includes an informative prior for the geometric mean of a lognormal distribution, defined by a point estimate and standard error (in the log scale). It was used in this project to define an informative prior based on the output of the Silica prediction model, and allows the combination of the prediction from the Silica database with new measurement data.

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 $f(\theta|x)$.

2 SEG.informedvar – prior

The joint prior distribution for the InformedVar model is given by

$$\begin{array}{rcl} \mu & \sim & U(\mu_0, \mu_1) \\ \log(\sigma) & \sim & N(\mu^*, \sigma^{*2}) \end{array} \tag{1}$$

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

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

The full conditional posterior density for μ is thus given by

$$\begin{aligned} f(\mu|\sigma, y) &\propto \exp\left\{-\frac{1}{2\sigma^2}\left(\sum y_i^2 - 2\mu\sum y_i + N\mu^2\right)\right\}I_\mu(\mu_0, \mu_1) \\ &\propto \exp\left\{-\frac{N}{2\sigma^2}\left(\mu^2 - 2\mu\bar{y}\right)\right\}I_\mu(\mu_0, \mu_1)\ , \end{aligned} \tag{3}$$

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

$$f(\sigma|\mu, y) \propto \frac{1}{\sigma^{N+1}} \exp\left\{-\frac{1}{2\sigma^2}\sum(y_i - \mu)^2\right\}\exp\left\{-\frac{(\log(\sigma) - \mu^*)^2}{2\sigma^{*2}}\right\} \tag{4}$$

Generating MCMC values for μ from its full conditional posterior density (3) is straightforward, while σ values will be sampled from its full conditional posterior density (4) 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.

2.1 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 p — measured without error — is given by

$$\begin{aligned} f(p|\mu, \sigma) &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(p_i - \mu)^2\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(p_i - \bar{p} + \bar{p} - \mu)^2\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[(p_i - \bar{p})^2 + 2(p_i - \bar{p})(\bar{p} - \mu) + (\bar{p} - \mu)^2]\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(p_i - \bar{p})^2\right\}\exp\left\{-\frac{n}{2\sigma^2}(\bar{p} - \mu)^2\right\} \\ &= \frac{1}{\sigma^n}\exp\left\{-\frac{(n-1)s_p^2}{2\sigma^2}\right\}\exp\left\{-\frac{n}{2\sigma^2}(\bar{p} - \mu)^2\right\}\ . \end{aligned} \tag{5}$$

The joint posterior for (μ, σ) is hence given by the product of (2) and the above likelihood of past data, that is,

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

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

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

2.1.1 Limitations / warnings

If it is thought that the past data were measured with error, they should 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.

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 scale as well in order to be usable.

3 Adding an informed prior on mean to SEG.informedvar [InformedMean]

In the original version of the algorithm, the prior distribution for μ is uniform on a specified range (μ_0, μ_1). In this section, we consider an alternative where prior information on the mean μ is included through a Normal prior distribution.

3.1 Context

Suppose that in a previous study pollution data Y^* was collected for a series of environmental conditions (indoors/outdoors premices, different concentrations of the potential contaminating product, different types of workers, etc.), fully described through a design matrix X and that the model $Y^* = X\beta + \epsilon$ was fit.

Suppose further that the scenario at stake in **our** study can be described through a design vector X_{pred} (where the variables in X_{pred} are the same as those in X); then our mean μ could be estimated through the use of the regression data (Y^*, X), with $\hat{\mu} = X'_{\text{pred}}\hat{\beta}$.

Remember that if c is a constant matrix (or vector) and θ a vector of random variables, then the variance of $c'\theta$ is given by

$$V(c'\theta) = c'V(\theta)c \ . \tag{6}$$

From linear regression theory, we know that the MLE for β is given by $\hat{\beta} = (X'X)^{-1}X'Y^*$; hence the variance of our μ estimate is given by

$$\begin{aligned}
V(\hat{\mu}) &= V(X'_{\text{pred}}\hat{\beta}) \\
&= X'_{\text{pred}}V(\hat{\beta})X_{\text{pred}} \quad \text{from (6)} \\
&= X'_{\text{pred}}V\left[(X'X)^{-1}X'Y^*\right]X_{\text{pred}} \\
&= X'_{\text{pred}}(X'X)^{-1}X'V(Y^*)X(X'X)^{-1}X_{\text{pred}} \quad \text{from (6)} \\
&= \hat{\sigma}^2 X'_{\text{pred}}(X'X)^{-1}X'X(X'X)^{-1}X_{\text{pred}} \quad \text{since } \ V(Y^*) = \hat{\sigma}^2 I \\
&= \hat{\sigma}^2 X'_{\text{pred}}(X'X)^{-1}X_{\text{pred}} \ .
\end{aligned}$$

Hence the mean of the estimate $\hat{\mu} = \mu$ (by construction) and its standard deviation is given by

$$\text{sd}(\hat{\mu}) = \hat{\sigma}\sqrt{X'_{\text{pred}}(X'X)^{-1}X_{\text{pred}}}$$

where $\hat{\sigma}$ is the residuals' sd from the regression model. The resulting point estimate $\hat{\mu}$ and its standard deviation (potentially slighly inflated to reflect an imperfect correspondance between the collected regression data and the currently collected data) can be used as moments, respectively labeled μ_r and σ_r, for the prior distribution for μ.

The regression data also provides information on σ but at this point it was decided to disregard it and stick to our own prior distribution on σ (through a logNormal distribution, as in the original version of `SEG.informedvar`).

3.2 Modifications to SEG.informedvar when using informed prior on mean

In the original model, the full conditional posterior distribution for μ is the product of two terms, where the first term (l_1, below) comes from the likelihood and the second is the prior distribution function for μ.

The first term of the likelihood, $l_1(\mu|\sigma, \ldots)$ is given by

$$l_1(\mu|\sigma, \ldots) = \begin{cases} \prod_i \exp\left\{-\frac{1}{2\sigma^2}(Y_i - \mu)^2\right\} & \text{when data is Normally distributed} \\ \prod_i \exp\left\{-\frac{1}{2\sigma^2}(\log(Y_i) - \mu)^2\right\} & \text{when data is log-Normally distributed} \end{cases}$$

depending on whether the outcome variable is normally or log-normally distributed, assuming that it is measured without error.

The second term of the the full conditional posterior distribution for μ, $f(\mu|\sigma,\ldots)$ is the prior for μ, which is $f(\mu) = I_\mu(\mu_0, \mu_1)$ in the original version of the algorithm.

The term $l_1(\mu|\sigma,\ldots)$ can be rewritten as

$$
\begin{aligned}
l_1(\mu|\sigma,\ldots) &= \prod_i \exp\left\{-\frac{1}{2}\lambda(\mu - z_i)^2\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(n\lambda\mu^2 - 2\mu\lambda\sum_i z_i\right)\right\} \qquad (7)
\end{aligned}
$$

where λ and the z_i's are defined differently depending on the data distribution.

Hence the full posterior conditional distribution for μ is given by

$$
\begin{aligned}
f(\mu|\sigma,\ldots) &= l_1(\mu|\sigma,\ldots)\cdot f(\mu) \\
&\propto \exp\left\{-\frac{1}{2}\left(n\lambda\mu^2 - 2\mu\lambda\sum_i z_i\right)\right\}\cdot I_\mu(\mu_0,\mu_1) \\
&= h(\mu; n\lambda, \lambda\sum_i z_i, \sigma)\cdot I_\mu(\mu_0,\mu_1) \ . \qquad (8)
\end{aligned}
$$

The algorithm used for sampling from h was described in earlier sections.

By adding an informed prior on μ as described in previous section, that is, by letting

$$
f(\mu) \propto \exp\left\{-\frac{1}{2\sigma_r^2}(\mu - \mu_r)^2\right\} \qquad (9)
$$

the full conditional posterior distribution for μ is changed to

$$
\begin{aligned}
f(\mu|\sigma,\ldots) &= l_1(\mu|\sigma,\ldots)\cdot f(\mu) \\
&\propto \exp\left\{-\frac{1}{2}\left(n\lambda\mu^2 - 2\mu\lambda\sum_i z_i\right)\right\} \quad \text{from (7)} \\
&\cdot \ \exp\left\{-\frac{1}{2\sigma_r^2}(\mu-\mu_r)^2\right\} \quad \text{from (9)} \\
&\propto \exp\left\{-\frac{1}{2}\left(n\lambda\mu^2 - 2\mu\lambda\sum_i z_i\right)\right\}\cdot\exp\left\{-\frac{1}{2\sigma_r^2}(\mu^2 - 2\mu\mu_r)\right\} \\
&= \exp\left\{-\frac{1}{2}\mu^2\left(n\lambda + \frac{1}{\sigma_r^2}\right)\right\}\cdot\exp\left\{-\frac{1}{2}\left(-2\mu\left[\lambda\sum_i z_i + \frac{\mu_r}{\sigma_r^2}\right]\right)\right\} \\
&= h(\mu; n\lambda + \frac{1}{\sigma_r^2}, \lambda\sum_i z_i + \frac{\mu_r}{\sigma_r^2}, \sigma) \quad \text{from (8)}
\end{aligned}
$$

that is, the shape of the full conditional posterior distribution $f(\mu|\sigma,\ldots)$ remains unchanged, it is just its parameters that are changed and its range which is no longer bounded to the interval (μ_0,μ_1). Hence the algorithm for sampling from μ's full conditional posterior distribution remains unchanged.

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 InformedVar and Two-Level InformedVar models.

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.