Respirable Crystalline Silica Exposure During Pavement Milling Using a Caterpillar Milling Machine Equipped with a Local Exhaust Ventilation System

Conducted with assistance from the Silica/Milling-Machines Partnership, affiliated with and coordinated through The National Asphalt Pavement Association (NAPA)

Duane R. Hammond, M.S., P.E.
Stanley A. Shulman, Ph.D.
Misty J. Hein, Ph.D.

Division of Applied Research and Technology
Engineering and Physical Hazards Branch
EPHB Report No. 282-26a

September 2014
Sites Surveyed:
Site 1: Hwy 371, Nisswa, Minnesota
Site 2: Hwy 53, south of Cook, MN
Site 3: Hwy 169, Ely, MN

NAICS Code: 237310 (Highway, Street, and Bridge Construction)

Survey Dates:
Site 1: July 23-25, 2013,
Site 2: August 27-29, 2013,
Site 3: September 23-25, 2013

Surveys Conducted By: Duane Hammond

Employer Representatives Contacted
Jim Domanus, Caterpillar
Bud Rife, Caterpillar
Tony Bodway (Chairman), Payne & Dolan
R. Gary Fore, National Asphalt Pavement Association (NAPA) consultant
Donald Elisburg, NAPA consultant

Analytical Work Performed by: Bureau Veritas North America
Disclaimer

Mention of any company or product does not constitute endorsement by NIOSH. In addition, citations to Web sites external to NIOSH do not constitute NIOSH endorsement of the sponsoring organizations or their programs or products. Furthermore, NIOSH is not responsible for the content of these Web sites. All Web addresses referenced in this document were accessible as of the publication date.
# Table of Contents

Abstract ............................................................................................... v
Introduction .......................................................................................... 1
  Background for Control Technology Studies ................................. 1
  Background for this Study ................................................................. 1
Occupational Exposure Limits and Health Effects ............................ 2
  Crystalline Silica Exposure Limits ..................................................... 3
Methodology ....................................................................................... 4
  Description of Evaluated Sites ......................................................... 5
Control Technology ............................................................................ 6
  Description of tested dust-emission control configuration ............. 6
Results ............................................................................................... 6
  Weather Observations ................................................................. 7
Discussion ......................................................................................... 8
Conclusions and Recommendations ................................................. 8
Acknowledgements ............................................................................ 9
References ....................................................................................... 9
Appendix: Statistical Analysis .......................................................... 18
  References: .................................................................................. 31
Abstract

Between July 23rd and September 25th, 2013, National Institute for Occupational Safety and Health (NIOSH) researchers and the Silica/Milling-Machines Partnership coordinated by the National Asphalt Pavement Association (NAPA) conducted field testing of a local exhaust ventilation (LEV) system on a Caterpillar PM200 cold milling machine. The tests included nine days of air sampling across three different highway construction sites in Minnesota. At each site, full-shift personal breathing zone samples for respirable crystalline silica were collected from the operator and ground man during normal work activities of asphalt pavement milling.

The data were analyzed two ways, (1) assuming the data were normally distributed and (2) assuming that they were lognormally distributed. For each distribution, a 97.5% upper confidence limit for the arithmetic mean respirable crystalline silica exposure for each occupation was calculated and compared to the NIOSH recommended exposure limit (REL) of 0.05 mg/m³. For results on either scale, 97.5% upper confidence limits are chosen for each occupation in order that the combined confidence is 95%. For the normal distribution analysis, the arithmetic mean respirable crystalline silica exposure for the operator was 0.052 mg/m³ with an upper 97.5% confidence limit of 0.071 mg/m³. The arithmetic mean respirable crystalline silica exposure for the ground man was 0.037 mg/m³ with an upper 97.5% confidence limit of 0.055 mg/m³. For the lognormal distribution analysis, the arithmetic mean respirable crystalline silica exposure for the operator was 0.052 mg/m³ with an upper 97.5% confidence limit of 0.083 mg/m³. The arithmetic mean respirable crystalline silica exposure for the ground man was 0.038 mg/m³ with an upper 97.5% confidence limit of 0.061 mg/m³. The 95% upper confidence limits for the arithmetic means for the operator and ground man are above the REL. For either analysis method it cannot be stated that the arithmetic mean exposures were below the REL for the population of sites from which those studied were chosen.

Based on the results of this study, NIOSH researchers recommend that Caterpillar consider refining their design to prevent clogging of the duct system before conducting additional field testing of the LEV dust controls. A possible solution to prevent clogging would be to further increase the open area at the intake to the LEV system so that the air intake velocity is lower without reducing the total volumetric flow-rate of air through the system. A lower intake air velocity should reduce the number of particles larger than the respirable size range of 10 µm from being drawn into the LEV system while keeping the drum housing and primary conveyor under negative pressure.

With these suggestions or other modifications to prevent clogging of the LEV dust controls, NIOSH researchers recommend that Caterpillar consider conducting additional field testing to verify that their final dust control design will reduce worker exposures below the NIOSH REL. The recommendations in this report are based on past successful dust control studies and would not prevent Caterpillar from pursuing other technologies, ideas, or inventions to reduce silica exposures on asphalt milling machines.
Introduction

Background for Control Technology Studies

The National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention (CDC) is the primary Federal agency engaged in occupational safety and health research. Located in the Department of Health and Human Services, it was established by the Occupational Safety and Health Act of 1970. This legislation mandated NIOSH to conduct a number of research and education programs separate from the standard setting and enforcement functions carried out by the Occupational Safety and Health Administration (OSHA) in the Department of Labor. An important area of NIOSH research deals with methods for controlling occupational exposure to potential chemical and physical hazards. The Engineering and Physical Hazards Branch (EPHB) of the Division of Applied Research and Technology has been given the lead within NIOSH to study the engineering aspects of health hazard prevention and control.

Background for this Study

NIOSH is studying the effectiveness of dust-emission controls during asphalt pavement-milling operations. Pavement-milling is the process of removing the road surface for recycling. The aim of this project is to determine if the dust emission-control systems installed on new pavement-milling machines and operated according to the manufacturers’ recommendations are adequate to control worker exposures below occupational exposure limits for respirable dust, especially that containing crystalline silica, a long-recognized occupational respiratory hazard. Chronic over-exposures to such dust may result in silicosis, a chronic progressive lung disease that eventually may be disabling or even fatal, and an increased risk of lung cancer [NIOSH 2002]. The long term goal of this project is to adequately control worker exposures to respirable dust and crystalline silica by providing data to support the development of best practice guidelines for engineering controls on asphalt pavement milling machines.

Many construction tasks have been associated with overexposure to crystalline silica [Rappaport et al. 2003]. Among these tasks are tuck pointing, concrete sawing, concrete grinding, and abrasive blasting [NIOSH 2000; Thorpe et al. 1999; Akbar-Khanzadeh and Brillhart 2002; Glindmeyer and Hammad 1988]. Road milling has also been shown to result in overexposures to respirable crystalline silica [Linch 2002; Rappaport et al. 2003; Valiante et al. 2004]. However, the three road-milling studies do not provide enough information about the operating parameters and engineering controls present on the milling machines to determine if the overexposures were due to a lack of effective controls or poor maintenance of the machines.

A variety of machinery are employed in asphalt pavement recycling, including cold-planers, heater-planers, cold-millers, and heater-scarifiers [Public Works 1995]. Cold-milling, which uses a toothed, rotating cutter drum to grind and remove the
pavement to be recycled, is primarily used to remove surface deterioration on both petroleum-asphalt aggregate and Portland-cement concrete road surfaces [Public Works 1995]. The milling machines used in cold-milling are the focus of this study.

The equipment evaluated during this study was a Caterpillar PM200 cold milling machine with a 2 m (79-inch) cutter drum and a diesel engine that provides 429 kilowatt (kW) (575 horsepower (hp)) at 1900 rpm. The Caterpillar PM200 was fitted with a local exhaust ventilation (LEV) system consisting of a hydraulic powered fan located on the secondary conveyor. The suction side of the fan was connected to two ducts each connected to a manifold that further split the flow and drew air from the drum housing and the conveyor transition area. The LEV system was designed to create a negative pressure in the drum housing and conveyor transition areas and to exhaust the air away from any worker locations.

This field study evaluated the performance of the LEV system using full-shift, time-weighted average personal breathing zone sampling for respirable dust and respirable crystalline silica exposures of the milling machine operator and ground man during three days at each of three sites. The study was conducted during the normal employee work activities on typical highway construction milling jobs.

This study was facilitated by the Silica/Milling-Machines Partnership, which is affiliated with and coordinated through the National Asphalt Pavement Association (NAPA). The partnership includes NAPA, the Association of Equipment Manufacturers (AEM), the manufacturers of almost all pavement-milling machines sold in the United States, numerous construction contractors, the International Union of Operating Engineers (IUOE), the Laborers’ International Union of North America (LIUNA), NIOSH, and other interested parties.

**Occupational Exposure Limits and Health Effects**

As a guide to the evaluation of hazardous workplace exposures, NIOSH investigators use mandatory and recommended occupational exposure limits (OELs) when evaluating chemical, physical, and biological agents in the workplace. Generally, OELs suggest exposure concentrations to which most workers may be exposed up to 10 hours per day, 40 hours per week for a working lifetime without experiencing adverse health effects. It is, however, important to note that not all workers will be protected from adverse health effects even though their exposures are maintained below these levels. A small percentage may experience adverse health effects because of individual susceptibility, a pre-existing medical condition, and/or hypersensitivity (allergy). In addition, some hazardous substances may act in combination with other workplace exposures, the general environment, or with medications or personal habits of the worker to produce health effects even if the occupational exposures are controlled at the level set by the exposure limit. Combined effects are determined using a mixture formula and are not considered in any individual OEL. Finally, OELs may change over the years as new information on the toxic effects of an agent become available.
Most OELs are expressed as a time-weighted average (TWA) exposure. A TWA exposure refers to the average airborne concentration of a substance during a normal 8- to 10-hour workday. In the United States, OELs have been established by federal agencies, professional organizations, state and local governments, and other entities. The U.S. Department of Labor OSHA permissible exposure limits (PELs) [29 CFR 1910.1000 (2003)] are occupational exposure limits that are legally enforceable in covered workplaces under the Occupational Safety and Health Act. NIOSH issues RELs that are based on a critical review of the scientific and technical information available on the prevalence of health effects, the existence of safety and health risks, and the adequacy of methods to identify and control hazards [NIOSH 1992]. They have been developed using a weight-of-evidence approach and formal peer review process. Other OELs that are commonly used and cited in the United States include the threshold limit values (TLVs®) recommended by American Conference of Governmental Industrial Hygienists (ACGIH®) [ACGIH 2013a]. ACGIH TLVs are considered voluntary guidelines for use by industrial hygienists and others trained in this discipline “to assist in the control of health hazards.” Workplace environmental exposure levels (WEELs®) are recommended OELs developed by the American Industrial Hygiene Association (AIHA®). WEELs have been established for some chemicals “when no other legal or authoritative limits exist” [AIHA 2007].

OSHA requires an employer to furnish employees a place of employment that is free from recognized hazards that are causing or are likely to cause death or serious physical harm (Occupational Safety and Health Act of 1970, Public Law 91–596, sec. 5(a)(1)). Thus, employers are required to comply with OSHA PELs. NIOSH investigators also encourage employers to consider other OELs in making risk assessment and risk management decisions to best protect the health of their employees. NIOSH investigators encourage the use of the traditional hierarchy-of-controls approach to eliminating or minimizing identified workplace hazards. This approach includes, in preferential order, the use of (1) substitution or elimination of the hazardous agent, (2) engineering controls (e.g., local exhaust ventilation, process enclosure) (3) administrative controls (e.g., limiting time of exposure, employee training, work practice changes, medical surveillance), and (4) personal protective equipment (e.g., respiratory protection, gloves, eye protection, hearing protection).

**Crystalline Silica Exposure Limits**

The NIOSH recommended exposure limit for respirable crystalline silica is 0.05 mg/m³ (50 µg/m³) as a TWA determined during a full-shift personal breathing zone sample. This REL is applicable for most workers who work up to a 10-hr workday during a 40-hr workweek to reduce the risk of developing silicosis, lung cancer, and other adverse health effects [NIOSH 2002]. In cases of simultaneous exposure to more than one form of crystalline silica, the concentration of free silica in air can be expressed (see equation 1) as micrograms of free silica per cubic meter of air sampled (µgSiO₂/m³) [NIOSH 1974],

\[
\text{Concentration of free silica} = \frac{\text{Micrograms of free silica}}{\text{Cubic meters of air}}
\]
\[
\frac{\mu gSiO_2}{m^3} = \frac{\mu gQ + \mu gC + \mu gT + \mu gP}{V}
\]  

(1)

where Q is quartz, C is cristobalite, T is tridymite, P is “other polymorphs,” and V is volume of air sampled in cubic meters.

The current OSHA PEL for respirable dust containing crystalline silica for the construction industry is measured by impinger sampling. In the construction industry, the PELs for cristobalite and quartz are the same. The PELs are expressed in millions of particles per cubic foot (mppcf) and calculated using the following formula [29 CFR 1926.55 (2003)]:

\[
\text{Respirable PEL} = \frac{250 \text{ mppcf}}{\% \text{Silica} + 5}
\]  

(2)

Since the PELs were adopted, the impinger sampling method has been rendered obsolete by gravimetric sampling [OSHA 1996]. OSHA currently instructs its compliance officers to apply a conversion factor of 0.1 mg/m³ per mppcf when converting between gravimetric sampling and the particle count standard when characterizing construction operation exposures [OSHA 2008].

On September 12, 2013, OSHA published a Notice of Proposed Rulemaking (NPRM) for occupational exposure to respirable crystalline silica. The NPRM is published in the Federal Register and proposes a PEL of 50 µg/m³ for respirable crystalline silica as an 8-hr TWA exposure [78 Fed. Reg. 56274 (2013)].

The ACGIH TLV for quartz and cristobalite (respirable fraction) is 0.025 mg/m³ [ACGIH 2013a]. The documentation to the TLV states that “it is the concern about fibrosis (silicosis) and the precedent inflammatory process resulting from silica exposures, and the association of inflammation and fibrosis with lung cancer that leads to this recommendation [ACGIH 2010].”

**Methodology**

NIOSH researchers conducted full-shift personal breathing zone sampling for respirable crystalline silica from the operator and ground man of a Caterpillar PM200 milling machine. The sampling was conducted over a total of nine days across three different highway construction sites during the course of normal employee work activities of milling asphalt pavement. The same machine was used for the entire study.

Personal breathing zone air samples for respirable dust and respirable crystalline silica were collected from the milling machine operator and ground man using respirable dust cyclones (model GK2.69, BGI Inc., Waltham, MA) at a flow rate of 4.2 liters/minute (L/min) with battery-operated sampling pumps (Gillian model
GilAir® Plus, Sensidyne®, Clearwater, FL) calibrated before and after each day’s use. A sampling pump was clipped to each sampled employee’s belt worn at their waist. The pump was connected via Tygon® tubing and a tapered Leur-type fitting to a pre-weighed, 37-mm diameter, 5-micron (μm) pore-size polyvinyl chloride filter supported by a backup pad in a three-piece filter cassette sealed with a cellulose shrink band (in accordance with NIOSH Methods 0600 and 7500) [NIOSH 1998, NIOSH 2003]. The front portion of the cassette was removed and the cassette was attached to a respirable dust cyclone.

The filter samples were analyzed for respirable particulates in accordance with NIOSH Method 0600 [NIOSH 1998]. The limit of detection (LOD) was 20 μg/sample. The limit of quantitation (LOQ) was 53 μg/sample. The results were blank corrected with the average of the media blanks.

Crystalline silica analysis of filter samples was performed using X-ray diffraction in accordance with NIOSH Method 7500 [NIOSH 2003]. The LODs for quartz, cristobalite and tridymite are 5 μg/sample, 10 μg/sample, and 10 μg/sample, respectively. The LOQs for quartz, cristobalite, and tridymite are 17 μg/sample, 33 μg/sample, and 33 μg/sample, respectively.

Bulk samples were analyzed in accordance with NIOSH Method 7500. The LODs for quartz, cristobalite, and tridymite in bulk samples are 0.3%, 0.3%, and 0.5%, respectively. The LOQs for quartz, cristobalite, and tridymite in bulk samples are 0.83%, 0.83%, and 1.7%, respectively.

**Description of Evaluated Sites**

The following is a description of the three highway construction sites included in the survey.

**Site 1: Hwy 371, Nisswa, Minnesota**

The milling machine removed between 4-inches and 6-inches of newer asphalt pavement using the full width of the drum during the first day of milling. The second and third day consisted of a full-depth removal of old asphalt pavement using the full width of the drum down to the sandy base material. Multiple other dust generating sources were present on the same construction site, including a broom machine and second asphalt milling machine. Dump truck traffic passed back and forth next to the milling operation all shift in addition to dump truck traffic serving the other milling machine on site.

**Site 2: Hwy 53, south of Cook, MN**

The milling machine removed between 6-inches and 8-inches of asphalt using the full width of the drum on the first day at the second site. The milling machine removed the asphalt shoulder during the second day at full-depth with a removal width that varied depending on the section of road but was typically about half the drum width. The third day consisted of full-width and full-depth asphalt milling. Multiple other dust generating sources were present on the same construction site.
including a broom machine, a motor grader, and another asphalt milling machine. Dump truck traffic passed back and forth next to the milling operation all shift in addition to dump truck traffic serving the other milling machine on site.

Site 3: Hwy 169, Ely, MN

The milling machine removed approximately 3-inches of asphalt pavement using the full width of the drum on the first day and the first three hours of the second day of milling. The last nine hours of the second day and all thirteen hours of the third day consisted of milling concrete curb and gutters with an occasional asphalt intersection. Multiple other dust generating sources were present on the same construction site including a second asphalt milling machine. A street sweeper was used instead of a broom machine. Dump truck traffic passed back and forth next to the milling operation all shift in addition to dump truck traffic serving the other milling machine on site. The local fire department was conducting an educational demonstration for the community at a distance of half a block from the milling operation during the evening of the third day of milling. The educational demonstration included burning debris and putting the fire out using various methods such as water or fire extinguishers. Visible smoke from the ongoing fire demonstration surrounded the vicinity of the milling operation for about one-hour.

Control Technology

Description of tested dust-emission control configuration

The equipment evaluated during this study was a Caterpillar PM200 cold milling machine with a 2 m (79-inch) cutter drum and a diesel engine that provides 429 kilowatt (kW) (575 horsepower (hp)) at 1900 rpm. The Caterpillar PM200 was fitted with an LEV system consisting of a hydraulic-powered fan located on the secondary conveyor. The suction side of the fan was connected to two ducts, each connected to a manifold that further split the flow and drew air from the drum housing and the conveyor transition area. The LEV system was designed to create a negative pressure in the drum housing and conveyor transition areas and to exhaust the air away from any worker locations.

Results

Full-shift personal breathing zone silica exposures during the nine days of sampling across three sites for the operator and ground man are shown in Table 1 along with the silica content in the bulk and filter samples for each day. Table 1 also shows the full-shift personal breathing zone respirable dust exposures for the operator and the ground man. At the sites studied, the percent bulk silica content varied between 14 and 59%, with an average of 31%.

The aim of this survey was to determine whether the engineering controls employed on this milling machine were able to control respirable silica exposures below the NIOSH REL of 0.05 mg/m³. This can be demonstrated if the upper
confidence limit for each occupation’s arithmetic mean is less than the REL. Since the data can be treated as either normal or lognormal distributions, confidence limits are provided for both scales.

PROC Mixed in SAS was the statistical computer package used for the analyses, for which estimates use the restricted maximum likelihood method [SAS Institute 2004]. A detailed explanation of the statistical method along with the SAS code used is provided in the Appendix. In brief, a single mixed-effect model combining data for the two occupations was used. Occupation was a fixed effect; site and days within sites were random effects. Thus, the data have three variance components: between sites, between days at sites, and residual (within days). For the results on either scale, 97.5% upper confidence limits were estimated for the arithmetic mean of each occupation to have an overall confidence of 95%.

The results on the original scale are shown in Table 2 and those from the log scale analysis are shown in Table 3. Residuals appeared to be normally distributed for the log scale analysis and marginally normally distributed for the original scale analysis (results not shown).

The between site variability for both scales was estimated to be 0, which simplified the analysis. Because the log of the geometric standard deviation (GSD) is comparable to the relative standard deviation (RSD), it makes sense to compare these quantities from Tables 2 and 3. Taking the natural log of the GSD estimates gives approximate RSD estimates. Thus, the day (at site) GSD for the log scale analysis of 1.515 corresponds to an RSD of about 42%, compared to 37.5% for the normal scale. The within day GSD of 1.415 from Table 3 corresponds to an RSD of about 35%, compared to 43.4% from Table 2. Thus, the precision is similar for the two scales.

For the normal distribution analysis, the arithmetic mean respirable crystalline silica exposure for the operator was 0.052 mg/m³ with an upper 97.5% confidence limit of 0.071 mg/m³. The arithmetic mean respirable crystalline silica exposure for the ground man was 0.037 mg/m³ with an upper 97.5% confidence limit of 0.055 mg/m³. For the lognormal distribution analysis, the arithmetic mean respirable crystalline silica exposure for the operator was 0.052 mg/m³ with an upper 97.5% confidence limit of 0.083 mg/m³. The arithmetic mean respirable crystalline silica exposure for the ground man was 0.038 mg/m³ with an upper 97.5% confidence limit of 0.061 mg/m³. Thus, from the results shown in Tables 2 and 3, the 95% upper confidence limits for the arithmetic means for the operator and ground man are each above the REL. Therefore, it cannot be stated that the arithmetic mean exposures were below the REL for the population of sites from which those studied were chosen.

Weather Observations

Table 4 through Table 6 show weather observations from the National Oceanic and Atmospheric Administration (NOAA) fixed weather station nearest to each evaluated site. Wind direction is reported as the angle, measured in a clockwise direction,
between true north and the direction from which the wind is blowing. Wind speed is reported as the rate of horizontal travel of air past a fixed point in meters per second (m/s). The average wind speed during the evaluation ranged from 0 to 7.7 m/s. Hourly temperatures during testing at the three evaluated sites ranged from 4°C (39°F) to 30°C (86°F). Relative humidity during testing at the three evaluated sites ranged from 41% to 100%.

**Discussion**

The LEV dust controls protected workers from exposures above the NIOSH REL during the first test site even though silica content in the recycled asphalt pavement was higher than NIOSH researchers have measured at any other evaluation site during the past ten years. However, partial clogging of the LEV dust controls resulted in approximately 40% reduced air velocity from 5,500 ft/min measured before the first day of testing to 3,300 ft/min by the end of the third day of testing at the first site. Some minimal cleaning of the LEV dust control system was performed between each day at the first two sites but not at the third site. Substantial cleaning of the LEV dust control system was performed between all test sites. Even with substantial cleaning in between sites, it was not possible to achieve air velocities comparable to the 5,500 ft/min measured before testing at the first site. Table 7 provides pre and post-shift centerline duct air velocity measurements. It was not always possible to collect air velocity measurements for all shifts or sites due to the work requirements of the crews and the inaccessibility of tools at certain highway locations. Air velocity measurements were not collected at the third site due to long work hours of the crew. The LEV dust control system appeared to be mostly clogged by the end of milling at the third site.

**Conclusions and Recommendations**

The evaluated Caterpillar PM200 needs modifications to prevent clogging of the LEV system so that worker exposure to respirable crystalline silica can be reduced during pavement milling operations. Measures to prevent clogging of the LEV dust control system would likely lead to reduced worker exposures to respirable crystalline silica.

Design changes should eliminate clogging in the LEV dust control system. It is important to design a smooth transition of air velocities between the duct and the take-off point of the dust control so that particles larger than the inhalable size range do not enter and clog the duct or damage the fan. The LEV dust control system could be designed to move air from the drum housing or primary conveyor through a hood or settling box to incrementally transition the air from a low velocity to a high velocity before air enters the duct. The design intent of the ventilation control should be to at least control respirable size silica particles in the 10 µm diameter and smaller size range. Particles less than 10 µm are capable of entering the gas exchange region of the lungs and triggering silicosis and cancer [Plog 2002]. One possible design would be to orient a hood or setting box so that air is pulled vertically from the drum housing or conveyor with a face velocity of
approximately 200 ft/min. Figure 1 shows a square settling box designed to settle out larger particles by transitioning air velocities from low to high before dust enters the duct.

With these suggestions or other modifications to prevent clogging of the LEV dust controls, NIOSH researchers recommend that Caterpillar consider conducting additional field testing to verify that the final dust control design will reduce worker exposures below the NIOSH REL. The recommendations in this report are based on past successful dust control studies [NIOSH 2013a, NIOSH 2013b]. These recommendations would not prevent Caterpillar from pursuing other technologies, ideas, or inventions to reduce silica exposures on asphalt milling machines.

**Acknowledgements**

The authors would like to thank management and staff at Reiner Contracting and Caterpillar Inc. for their gracious hospitality and assistance during this testing. Their commitment to the development and testing of engineering controls to reduce occupational exposures to respirable crystalline silica is important to protect worker safety and health.

**References**

ACGIH® [2010]. 2010 TLVs® and BEIs®: threshold limit values for chemical substances and physical agents and biological exposure indices. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.


NIOSH [2013b]. Control technology for crystalline silica exposure during pavement milling using a Wirtgen milling machine equipped with a vacuum cutting system. By Hammond DR, Shulman SA. Cincinnati, OH: U.S. Department of Health and Human


R. Project [2012]. The R project for statistical computing [http://www.r-project.org/].


Table 1: Full-shift personal breathing zone respirable dust and silica exposures for the operator and ground man in mg/m³

<table>
<thead>
<tr>
<th>Site 1 Day 1</th>
<th>Site 1 Day 2</th>
<th>Site 1 Day 3</th>
<th>Site 2 Day 4</th>
<th>Site 2 Day 5</th>
<th>Site 2 Day 6</th>
<th>Site 3 Day 7</th>
<th>Site 3 Day 8</th>
<th>Site 3 Day 9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Respirable Dust (mg/m³)</strong></td>
<td>23-Jul</td>
<td>24-Jul</td>
<td>25-Jul</td>
<td>27-Aug</td>
<td>28-Aug</td>
<td>29-Aug</td>
<td>23-Sep</td>
<td>24-Sep</td>
</tr>
<tr>
<td>Operator</td>
<td>0.213</td>
<td>0.194</td>
<td>0.124</td>
<td>0.119</td>
<td>0.215</td>
<td>0.388</td>
<td>0.201</td>
<td>0.435</td>
</tr>
<tr>
<td>Groundman</td>
<td>0.225</td>
<td>0.158</td>
<td>0.099</td>
<td>0.122</td>
<td>0.353</td>
<td>0.310</td>
<td>0.192</td>
<td>0.177</td>
</tr>
<tr>
<td><strong>Respirable Crystalline Silica (mg/m³)</strong></td>
<td>23-Jul</td>
<td>24-Jul</td>
<td>25-Jul</td>
<td>27-Aug</td>
<td>28-Aug</td>
<td>29-Aug</td>
<td>23-Sep</td>
<td>24-Sep</td>
</tr>
<tr>
<td>Operator</td>
<td>0.047</td>
<td>0.041</td>
<td>0.020</td>
<td>0.021</td>
<td>0.042</td>
<td>0.069</td>
<td>0.031</td>
<td>0.084</td>
</tr>
<tr>
<td>Groundman</td>
<td>0.041</td>
<td>0.029</td>
<td>0.016</td>
<td>0.021</td>
<td>0.067</td>
<td>0.043</td>
<td>0.029</td>
<td>0.023</td>
</tr>
<tr>
<td>% silica bulk</td>
<td>30%</td>
<td>44%</td>
<td>57%*</td>
<td>27%</td>
<td>34%</td>
<td>33%</td>
<td>45%</td>
<td>16%**</td>
</tr>
<tr>
<td></td>
<td>29%</td>
<td>37%</td>
<td>39%</td>
<td>37%</td>
<td>42%</td>
<td>59%</td>
<td>15%**</td>
<td>14%**</td>
</tr>
<tr>
<td></td>
<td>34%</td>
<td>37%</td>
<td>30%</td>
<td>30%</td>
<td>32%</td>
<td>33%</td>
<td>18%**</td>
<td>22%**</td>
</tr>
<tr>
<td>% silica operator</td>
<td>22%</td>
<td>21%</td>
<td>16%</td>
<td>18%</td>
<td>20%</td>
<td>18%</td>
<td>15%</td>
<td>19%</td>
</tr>
<tr>
<td>% silica groundman</td>
<td>18%</td>
<td>18%</td>
<td>16%</td>
<td>17%</td>
<td>19%</td>
<td>14%</td>
<td>15%</td>
<td>13%</td>
</tr>
<tr>
<td>Sample time operator (min)</td>
<td>481</td>
<td>601</td>
<td>672</td>
<td>661</td>
<td>678</td>
<td>276</td>
<td>701</td>
<td>714</td>
</tr>
<tr>
<td>Sample time groundman (min)</td>
<td>477</td>
<td>602</td>
<td>672</td>
<td>665</td>
<td>676</td>
<td>276</td>
<td>694</td>
<td>715</td>
</tr>
</tbody>
</table>

*The 57% silica bulk sample on site 1, day 3 was from the sand below the asphalt layer.

**All Day 8 and Day 9 bulk samples were from milling concrete.
### Table 2: Data treated as normally distributed with estimates from the application of SAS PROC MIXED using restricted maximum likelihood*

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Arithmetic mean (AM), mg/m³, from models</th>
<th>Relative standard deviation (RSD), between sites</th>
<th>Relative standard deviation (RSD), for days at sites, relative to average of arithmetic means</th>
<th>Relative standard deviation (RSD), within days, relative to average of arithmetic means</th>
<th>Simultaneous Upper 95% Confidence Limits (AM) for both Occupations, mg/m³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground man</td>
<td>0.037</td>
<td>0</td>
<td>0.375</td>
<td>0.434</td>
<td>0.055</td>
</tr>
<tr>
<td>Operator</td>
<td>0.052</td>
<td>0</td>
<td>0.071</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*PROC MIXED is an application in SAS [SAS 2004]. Estimates use restricted maximum likelihood. Estimates are obtained from a single model for the two occupations. Upper confidence limits were obtained by multiplying the value of the 97.5th percentile of the t distribution with 13.5 degrees of freedom by the standard error of the estimate, which is added to the sample average. The fractional degrees of freedom are from Satterthwaite calculations. [SAS 2004] Overall 95% confidence is obtained for both occupations.

### Table 3: Data treated as lognormally distributed with estimates from the application of SAS PROC MIXED using restricted maximum likelihood*

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Arithmetic mean (AM), mg/m³, from model</th>
<th>Geometric mean (GM), mg/m³</th>
<th>Geometric standard deviation (GSD), between sites</th>
<th>Geometric standard deviation (GSD), between days at sites</th>
<th>Geometric standard deviation (GSD), within days</th>
<th>Simultaneous Upper 95% Confidence Limits (AM) for both Occupations, mg/m³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground man</td>
<td>0.038</td>
<td>0.033</td>
<td>1.0, which means Log scale GSD=0</td>
<td>1.515</td>
<td>1.415</td>
<td>0.061</td>
</tr>
<tr>
<td>Operator</td>
<td>0.052</td>
<td>0.045</td>
<td></td>
<td>3</td>
<td></td>
<td>0.083</td>
</tr>
</tbody>
</table>

*PROC MIXED is an application in SAS [SAS 2004]. Estimates are obtained from a single model for the two occupations. As for the normally distributed results shown above, the confidence limits are based on the t distribution, but here include the variances associated with the geometric mean and the two non-zero variance components. For small sample sizes of lognormal data, the nominal confidence for the calculated confidence interval is not the actual confidence. Simulations were used to determine that use of 98.5% upper confidence limits gives approximately 97.5% confidence individually for each occupation. Overall 95% confidence is obtained for both occupations. Confidence limits computed on the log scale were exponentiated. 11.9 degrees of freedom are used in computations, based on Satterthwaite calculations.[SAS 2004]
Table 4: Wind speed and direction near Site 1 from July 23-25 on Hwy 371 N (Nisswa, MN)

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>653</th>
<th>753</th>
<th>853</th>
<th>953</th>
<th>1053</th>
<th>1153</th>
<th>1253</th>
<th>1353</th>
<th>1453</th>
<th>1553</th>
<th>1653</th>
<th>1753</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>330</td>
<td>330</td>
<td>350</td>
<td>340</td>
<td>30</td>
<td>20</td>
<td>30</td>
<td>30</td>
<td>10</td>
<td>10</td>
<td>360</td>
<td>360</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>3.6</td>
<td>4.1</td>
<td>3.1</td>
<td>2.1</td>
<td>3.1</td>
<td>3.6</td>
<td>3.1</td>
<td>3.6</td>
<td>3.6</td>
<td>2.6</td>
<td>3.1</td>
<td>4.6</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>16.1</td>
<td>15.6</td>
<td>14.4</td>
<td>13.9</td>
<td>14.4</td>
<td>15.6</td>
<td>16.1</td>
<td>17.8</td>
<td>17.8</td>
<td>19.4</td>
<td>20.6</td>
<td>21.7</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>75</td>
<td>78</td>
<td>84</td>
<td>87</td>
<td>84</td>
<td>80</td>
<td>70</td>
<td>60</td>
<td>63</td>
<td>55</td>
<td>53</td>
<td>47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>653</th>
<th>753</th>
<th>853</th>
<th>953</th>
<th>1053</th>
<th>1153</th>
<th>1253</th>
<th>1353</th>
<th>1453</th>
<th>1553</th>
<th>1653</th>
<th>1753</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>330</td>
<td>220</td>
<td>250</td>
<td>260</td>
<td>-</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>2.1</td>
<td>2.6</td>
<td>2.1</td>
<td>2.1</td>
<td>2.6</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>13.9</td>
<td>13.3</td>
<td>15</td>
<td>15</td>
<td>14.4</td>
<td>16.1</td>
<td>18.3</td>
<td>19.4</td>
<td>21.1</td>
<td>22.2</td>
<td>21.7</td>
<td>23.3</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>89</td>
<td>93</td>
<td>87</td>
<td>87</td>
<td>90</td>
<td>87</td>
<td>81</td>
<td>68</td>
<td>59</td>
<td>55</td>
<td>55</td>
<td>46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>653</th>
<th>753</th>
<th>853</th>
<th>953</th>
<th>1053</th>
<th>1153</th>
<th>1253</th>
<th>1353</th>
<th>1453</th>
<th>1553</th>
<th>1653</th>
<th>1753</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>220</td>
<td>210</td>
<td>210</td>
<td>180</td>
<td>210</td>
<td>310</td>
<td>310</td>
<td>350</td>
<td>300</td>
<td>340</td>
<td>-</td>
<td>300</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>3.1</td>
<td>2.6</td>
<td>2.1</td>
<td>1.5</td>
<td>2.1</td>
<td>1.5</td>
<td>2.6</td>
<td>1.5</td>
<td>2.1</td>
<td>3.1</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>17.8</td>
<td>17.2</td>
<td>16.1</td>
<td>15</td>
<td>14.4</td>
<td>17.8</td>
<td>18.9</td>
<td>20.6</td>
<td>21.7</td>
<td>22.8</td>
<td>24.4</td>
<td>25.6</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>84</td>
<td>87</td>
<td>90</td>
<td>90</td>
<td>93</td>
<td>87</td>
<td>84</td>
<td>73</td>
<td>71</td>
<td>64</td>
<td>56</td>
<td>52</td>
</tr>
</tbody>
</table>
Table 5: Wind speed and direction near Site 2 on Hwy 53 on August 27-29 (Cook, MN)

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>0657</th>
<th>0757</th>
<th>0857</th>
<th>0957</th>
<th>1056</th>
<th>1156</th>
<th>1256</th>
<th>1356</th>
<th>1456</th>
<th>1556</th>
<th>1656</th>
<th>1757</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>20</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>20</td>
<td>22</td>
<td>26</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>94</td>
<td>94</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>74</td>
<td>70</td>
<td>66</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>656</th>
<th>756</th>
<th>856</th>
<th>950</th>
<th>1056</th>
<th>1157</th>
<th>1256</th>
<th>1456</th>
<th>1556</th>
<th>1656</th>
<th>1756</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>260</td>
<td>240</td>
<td></td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>16</td>
<td>15</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>24</td>
<td>28</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>94</td>
<td>94</td>
<td>100</td>
<td>94</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>83</td>
<td>70</td>
<td>59</td>
<td>55</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>657</th>
<th>756</th>
<th>856</th>
<th>956</th>
<th>1056</th>
<th>1156</th>
<th>1256</th>
<th>1356</th>
<th>1457</th>
<th>1556</th>
<th>1657</th>
<th>1756</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>80</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>140</td>
<td>250</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>3.1</td>
<td>2.6</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>18</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>18</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>88</td>
<td>88</td>
<td>89</td>
<td>94</td>
<td>94</td>
<td></td>
</tr>
</tbody>
</table>
Table 6: Wind speed and direction near site 3 September 23-25 on Hwy 169 (Ely, MN)

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>0653</th>
<th>0753</th>
<th>0853</th>
<th>0953</th>
<th>1053</th>
<th>1153</th>
<th>1253</th>
<th>1353</th>
<th>1453</th>
<th>1553</th>
<th>1653</th>
<th>1753</th>
<th>1853</th>
<th>1953</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>130</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>150</td>
<td>150</td>
<td>170</td>
<td>160</td>
<td>170</td>
<td>160</td>
<td>170</td>
<td>160</td>
<td>170</td>
<td>170</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>1.5</td>
<td>2.1</td>
<td>3.6</td>
<td>3.1</td>
<td>3.6</td>
<td>2.6</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
<td>5.7</td>
<td>5.7</td>
<td>7.2</td>
<td>7.7</td>
<td>6.2</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>14</td>
<td>16</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>19</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>82</td>
<td>71</td>
<td>67</td>
<td>63</td>
<td>56</td>
<td>52</td>
<td>49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>0652</th>
<th>0751</th>
<th>0853</th>
<th>0953</th>
<th>1053</th>
<th>1153</th>
<th>1253</th>
<th>1353</th>
<th>1453</th>
<th>1553</th>
<th>1653</th>
<th>1753</th>
<th>1853</th>
<th>1953</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>180</td>
<td>180</td>
<td>200</td>
<td>180</td>
<td>170</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>190</td>
<td>200</td>
<td>180</td>
<td>190</td>
<td>190</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>3.1</td>
<td>4.1</td>
<td>3.1</td>
<td>2.1</td>
<td>2.6</td>
<td>2.1</td>
<td>2.6</td>
<td>3.1</td>
<td>4.1</td>
<td>4.1</td>
<td>3.1</td>
<td>2.1</td>
<td>4.6</td>
<td>4.1</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>15</td>
<td>17</td>
<td>19</td>
<td>21</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>87</td>
<td>82</td>
<td>82</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>82</td>
<td>82</td>
<td>72</td>
<td>63</td>
<td>60</td>
<td>49</td>
<td>46</td>
<td>43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (HrMn)</th>
<th>0652</th>
<th>0751</th>
<th>0853</th>
<th>0953</th>
<th>1053</th>
<th>1153</th>
<th>1253</th>
<th>1353</th>
<th>1453</th>
<th>1553</th>
<th>1653</th>
<th>1753</th>
<th>1852</th>
<th>1952</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind direction (°)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>160</td>
<td>160</td>
<td>170</td>
</tr>
<tr>
<td>Wind Speed (m/s)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>2.6</td>
<td>3.6</td>
<td>3.1</td>
<td>5.1</td>
<td>4.1</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>11</td>
<td>16</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Relative Humidity %</td>
<td>99</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>87</td>
<td>72</td>
<td>56</td>
<td>49</td>
<td>46</td>
<td>41</td>
<td>41</td>
<td>41</td>
</tr>
</tbody>
</table>
Table 7: Pre and post-shift centerline duct velocity measurements in ft/min

<table>
<thead>
<tr>
<th></th>
<th>Site 1 Day 1</th>
<th>Site 2 Day 4</th>
<th>Site 2 Day 5</th>
<th>Site 2 Day 6</th>
<th>Site 3 Day 7</th>
<th>Site 3 Day 8</th>
<th>Site 3 Day 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-shift</td>
<td>5,500</td>
<td>4,800</td>
<td>4,000</td>
<td>2,400</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>post-shift</td>
<td>5,300</td>
<td>3,300</td>
<td>3,000</td>
<td>2,000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Air velocity measurements were not collected for all shifts due to the work requirements of the crews and the inaccessibility of tools at certain highway locations. The LEV dust control system was cleaned out between days and between sites during testing at sites 1 and 2.

**Air velocity measurements were not collected during site 3 due to long work hours of the crew. The LEV dust control system was cleaned out before testing at site 3 but was not cleaned out between days at site 3.

Figure 1: Hood with settling box on a milling machine
Appendix: Statistical Analysis

The nine measurements each for the operator and ground man were used in one statistical model, allowing for means for the two occupations and variance component estimates for sites, days at sites, and the residual (within days), using restricted maximum likelihood (REML) estimates with Proc Mixed [SAS 2004]. Models were fitted both on the original scale and on the natural log transformed data. The log scale model used is:

\[ Y_{sdj} = \beta_0 + \beta_1 (\text{Job}=\text{OP}) + a_s + b_{d(s)} + e_{sdj} \quad (A-1) \]

where

- \( Y_{sdj} \) = the natural log transformed measurement at site s, day d at site s, and job j,
- \( a_s \) = random, mean 0 normally distributed value for the site s with variance \( \sigma^2_s \),
- \( b_{d(s)} \) = random, mean 0 normally distributed value for day d at site s with variance \( \sigma^2_{d(s)} \),
- \( e_{sdj} \) = random, mean 0 normally distributed value for job j at site s on day d with variance \( \sigma^2_r \),
- \( \beta_0 \) = the ground man’s log scale mean,
- \( \beta_0 + \beta_1 \) = and the operator’s log scale mean,
- \( \text{(Job}=\text{OP}) = 1 \) for operator, 0 for ground man.

The original scale model is written analogously, but uses the exponential of \( Y_{sdj} \). The aim of the data analysis is to place an upper confidence limit on the arithmetic mean of the exposures for each job. This is done very easily when the data are analyzed on the original scale. However, because exposure data are usually assumed to be log-normally distributed, the data must be analyzed on the log scale. The statistical procedures are then more complicated. We do analyses on each scale for completeness.

For each occupation, for both the normal and lognormal approaches, the 97.5 percent upper confidence limits are calculated, in order to have 95% confidence overall for both jobs.
Modeling Issues

For the model using either the ln scale or the original scale data, the estimated site variance (based on the REML estimation method) was 0. In the Mixed procedure, variance components can be constrained to be non-negative. A second model for the ln scale data that fits separate site variances for the ground man and operator (by specifying the “Group=job” option in the “random” statement for “sites” or by including an interaction term between sites and jobs in the model) produced estimated variances of 0 and 0.054, respectively. Results based on F tests using the model mean squares from the sum of squares estimation method can be used to test whether the site variance components are needed in the model [Littell 2006]. For the ln scale analysis, the two F tests (one testing the mean square for sites and the other testing the mean square for job by site interaction) had p-values greater than 0.25. By contrast, the days at sites mean square has a p-value of 0.036. For the original scale analysis, the p-values exceed 0.19 for these two F tests. By contrast the days at sites mean square has a p-value of 0.087. These results indicate that the variance components for site and for job by site are not needed in the models on either scale.

Upper confidence limits on the normal scale

The confidence limits for each occupation mean on the normal scale use the standard formula:

\[
UCL = \bar{x} + t_{13.5,0.975} \left( \frac{s}{\sqrt{\bar{x}}} \right) \tag{A-2}
\]

where

- \(UCL\) = upper confidence limit,
- \(\bar{x}\) = occupational mean,
- \(t_{13.5,0.975}\) = the 97.5\textsuperscript{th} percentile of the Student’s t distribution with 13.5 degrees of freedom,
- \(s/\sqrt{\bar{x}}\) = standard error of occupation mean estimate.

The degrees of freedom, obtained by Satterthwaite approximation [SAS 2004], are fractional because the standard error is a linear combination of the between days and the residual mean squares. The upper confidence limits obtained by this method are 0.055 mg/m\textsuperscript{3} for the ground man and 0.071 mg/m\textsuperscript{3} for the operator, both of which are >0.05 mg/m\textsuperscript{3}.
Upper confidence limits on the lognormal scale

For the lognormal scale analysis, confidence limits for the occupation arithmetic means are obtained as follows. The natural log of the arithmetic mean for each job is:

$$\text{Ln}(\mu_{am,j}) = \log_{e}(\mu_{j}) + 0.5 \cdot (\sigma^2_s + \sigma^2_d(s) + \sigma^2_r)$$ (A-3)

where

$$\text{Ln}(\mu_{am,j}) = \text{The natural log of the arithmetic mean for job } j,$$
$$\mu_{j} = \log_{e}(\text{job mean; natural log of geometric mean for job } j),$$
$$\sigma^2_s = \text{between sites variance},$$
$$\sigma^2_d(s) = \text{days at sites variance},$$
$$\sigma^2_r = \text{residual variance}.$$

where all parameters in the above expression are on the natural log scale. Thus, in the above expression, “job mean” refers to the natural log of the geometric mean. Let $E_1$ denote the quantity obtained by replacing the parameters in the above expression by sample estimates:

$$E_1 = \hat{\mu}_{j} + 0.5 \cdot (\hat{\sigma}^2_s + \hat{\sigma}^2_d(s) + \hat{\sigma}^2_r),$$

Where $\hat{\mu}_{j}$ is an estimate of $\mu_{j}$, and the “hat” notation is used for the other quantities.

The variance of $E_1$ is required to estimate the upper confidence limit. The Mixed procedure provides the standard errors of the occupation mean estimates, and also gives estimated asymptotic variances and covariances for the variance estimates, which are used in the computations. For balanced data the occupation mean estimate is independent of the variance estimates. The variance of $(E_1)$ can be estimated by using the variance estimates produced by the Mixed procedure:

$$\text{Var}(E_1) = \hat{\sigma}^2_{\mu_j} + 0.25 \cdot [\text{Var}(\hat{\mu}_{j}) + \text{Var}(\hat{\sigma}^2_s) + \text{Var}(\hat{\sigma}^2_d(s)) + \text{Var}(\hat{\sigma}^2_r)]$$ (A-4)

$$+ 0.5 \cdot [c_{\text{cov}}(\hat{\sigma}^2_s, \hat{\sigma}^2_d(s)) + c_{\text{cov}}(\hat{\sigma}^2_s, \hat{\sigma}^2_r) + c_{\text{cov}}(\hat{\sigma}^2_d(s), \hat{\sigma}^2_r)]$$

In the above expression $\hat{\sigma}^2_{\mu_j}$ denotes the squared standard error of the Ln scale job mean, and each $\text{Var}$ denotes an estimated asymptotic variance and each $c_{\text{cov}}$ denotes an estimated asymptotic covariance. Formula A-4 can be derived either by using the formula for the variance of a sum or from the delta method [Mandel 2013].
In Mandel’s notation:
\[ g(\theta) = g(\mu_j, \sigma_s^2, \sigma_{d(e)}^2, \sigma_e^2) = \mu_j + 0.5 \times (\sigma_s^2 + \sigma_{d(e)}^2 + \sigma_e^2), \]
and the partial derivatives are:
\[
\begin{align*}
\operatorname{transpose}(D(\theta)) &= \left( \frac{\partial g(\theta)}{\partial \mu_j}, \frac{\partial g(\theta)}{\partial \sigma_s^2}, \frac{\partial g(\theta)}{\partial \sigma_{d(e)}^2}, \frac{\partial g(\theta)}{\partial \sigma_e^2} \right) \\
(1, 0.5, 0.5, 0.5).
\end{align*}
\]

Finally, the required upper confidence limits are produced using the delta method:
\[
\text{UCL} = \exp\left[ E1 + \text{t}_\text{stat} \times \sqrt{\text{Var}(E1)} \right]
\]

where “t_stat” is the appropriate percentile of the t-distribution. The difference from the presentation in [Mandel 2013] is that the reference uses the normal rather than t percentiles, which are used here because of the small sample sizes.

An important issue involves selection of appropriate t-distribution critical values which requires specification of both degrees of freedom and a confidence level. Because actual confidence interval coverage may be less than nominal coverage when sample sizes are small, simulations were used to determine the percentile of the t distribution to use as the “t_stat” values in Equation A-5 (see SAS simulation code below in program 2). The degrees of freedom used for the critical value is the degrees of freedom associated with the estimated job mean.

The simulations involve generating log-normally distributed data with geometric means for the ground man and operator jobs similar to the observed data and realistic values for the site, days within site, and residual variances. Since the true values of the variances are not known, the SAS program evaluates different scenarios that are specified by the user (i.e., different true values for the job means and the variances). The actual data have a total geometric standard deviation of about 1.7. In the simulations, the variances are chosen so that the total geometric standard deviation varies between 1.55 and 2.1. Two choices of the between site variances are used in the simulations: 0, as estimated from the model, and 0.05. The latter is included in order to allow for the possibility that there may be nonzero between site variance. In general, the user must specify a range of values over which to carry out simulations. The percentile of the t distribution that achieves approximately 97.5% confidence over this range should be used.

The results of the simulations relating to the Caterpillar data are shown in Table A-1 below, for between site variances of 0 and 0.05. For a between site variance of 0, use of the 98.5th percentile for the t distribution leads to confidence greater than 98%, whereas use of the 97.5th percentile leads to confidence < 97.5% for several scenarios. However, the results for a between sites variance of 0.05 suggest that
the 98.5\textsuperscript{th} percentile is required. The confidence levels are only approximate. There is judgment involved in selecting the range of values to evaluate the coverage.

The resulting upper confidence limits, 0.061 mg/m\textsuperscript{3} for the ground man and 0.083 mg/m\textsuperscript{3} for the operator, indicate that neither occupation’s mean has an upper confidence limit less than 0.05 mg/m\textsuperscript{3}, at the overall 95% confidence level. The 98.5\textsuperscript{th} percentile of the t distribution with 11.9 degrees of freedom is used in the computations, based on the Satterthwaite approximation [SAS 2004]. See SAS program 1 below.

The SAS simulation code in program 2 also computes the fraction of simulations (the probability) for which the upper confidence limit for each model is less than the REL, and also computes the median and mean of the upper confidence limits. By changing the number of sites, we can see the benefit of increasing the number of sites for the ground man and operator evaluations. The ground man data have the lower geometric mean. However for that job when the parameter estimates are used as the true values in the simulations, even for five sites, three days per site, this fraction (probability) is less than 0.40. This suggests that the control system requires improvement if it is to demonstrate that it can meet the REL with an economically feasible number of runs.

The above use of the simulations indicates how they can be used to determine if inclusion of an extra site can likely demonstrate that the REL is met, when the results from the first three sites almost meet the REL. Following this procedure will reduce the overall confidence but can be beneficial if can be demonstrated that the control system meets the REL.

Table A-1. Coverage of upper confidence limits using nominal 97.5\textsuperscript{th} and 98.5\textsuperscript{th} t\textsubscript{stat} percentiles, for various choices of variance component parameters, based on 5000 simulations per scenario.

<table>
<thead>
<tr>
<th>Scenario:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>These are estimates from the data:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td>var(site) = 0</td>
<td></td>
</tr>
<tr>
<td>var(day) = 0.17</td>
<td>var(day) = 0.20</td>
<td>var(res) = 0.12</td>
<td>var(res) = 0.30</td>
<td>var(res) = 0.20</td>
<td>var(res) = 0.30</td>
<td>var(res) = 0.20</td>
<td>var(res) = 0.30</td>
<td>var(res) = 0.20</td>
<td>var(res) = 0.30</td>
<td></td>
</tr>
<tr>
<td>var(res)* = 0.12</td>
<td>var(res) = 0.15</td>
<td>var(res) = 0.20</td>
<td>var(res) = 0.30</td>
<td>var(res) = 0.20</td>
<td>var(res) = 0.30</td>
<td>var(res) = 0.20</td>
<td>var(res) = 0.30</td>
<td>var(res) = 0.20</td>
<td>var(res) = 0.30</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GMs from data:</th>
<th>Nominal confidence level of t\textsubscript{stat}</th>
<th>In(GM)</th>
<th>ln(GM)</th>
<th>ln(GM)</th>
<th>ln(GM)</th>
<th>ln(GM)</th>
<th>ln(GM)</th>
<th>ln(GM)</th>
<th>ln(GM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In(GM) = 3.11</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
</tr>
<tr>
<td>ln(GM) = 3.41</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
</tr>
<tr>
<td>In(GM) = 3.11</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
</tr>
<tr>
<td>ln(GM) = 3.41</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
<td>97.5</td>
</tr>
</tbody>
</table>
*Var(res) is the variance of the residual

Another important issue concerns the use of the asymptotic variance-covariance matrix from the Mixed procedure in SAS to estimate the variance of \( E_1 \). By far, the largest portion of the variance of \( E_1 \) is due to the squared standard error of the job mean. To see this, consider the sum of squares estimates, the method based on mean squares, which are identical to the REML estimates when the data design is balanced. For the model fitted with a site variance equal to 0, it can be shown that the squared standard error of each occupation mean is estimated by \( (\text{MS}_{\text{days}} + \text{MS}_{\text{residual}})/(18) \), where \( \text{MS}_{\text{days}} \) and \( \text{MS}_{\text{residual}} \) denote, respectively, the mean squares for days at sites and for the residual. Furthermore, the variance of \( 0.5 \times (\text{days at site variance estimate} + \text{residual variance estimate}) \) can be shown to be equal to \( \left[ \text{var}(\text{MS}_{\text{days}}) + \text{var}(\text{MS}_{\text{residual}}) \right]/16 \). For the chi square distribution, the variance of a mean square is twice the square of the expectation of the mean square, divided by the degrees of freedom (=8 for each mean squares). The standard error of the job mean was the dominant component of the standard error of \( E_1 \) for the variance values considered here.

Similarly, for the three variance component model, the squared estimated standard error of each job mean can be shown to be \( (\text{MS}_{\text{sites}} + \text{MS}_{\text{residual}})/18 \), and the variance of \( 0.5 \times (\text{sites variance estimate} + \text{days at site variance estimate} + \text{residual variance estimate}) \) can be shown to be \( \left[ \text{var}(\text{MS}_{\text{sites}}) + 4 \times \text{var}(\text{MS}_{\text{days}}) + 9 \times \text{var}(\text{MS}_{\text{residual}}) \right]/144 \), where \( \text{MS}_{\text{sites}} \) denotes the mean square for sites. As for the two component model, for relatively small variance component values (much less than 1), the standard error will be the largest portion of the standard error of \( E_1 \).

For the values considered in the simulation, the standard error of the job mean was at least 90% of the standard error of \( E_1 \), based on the above calculations. There are two consequences of this:

1) The degrees of freedom for \( E_1 \) should be close to that of the standard error for the job mean.

2) The asymptotic variance-covariance matrix that the Mixed procedure provides is obtained by substituting sample mean square estimates for the expectations of the mean squares. These are biased estimates (see [Searle 1992] for the unbiased estimates); however, because the variance associated with the sample variances is such a small portion of the total standard deviation, there is no problem using the asymptotic variance-covariance matrix from the MIXED procedure. The standard errors of \( E_1 \) which use the asymptotic variances are between 7 and 15% larger than those based on the unbiased estimates. Whereas the 98.5% probability values of t_stat suffice when the asymptotic variances are used, the 99% probability values are required when the unbiased estimates are used, because of the smaller standard errors associated with the unbiased estimates. Thus, for either choice of variance estimator, appropriate selection of the t_stat value, based on the simulations, will provide the required
coverage, adjusting both for deviations from normality of $E_1$, and also for the bias in the estimation of the variance–covariance matrix of the variance estimates, when the asymptotic variance estimates are used.

It is useful to apply the same methodology when the data come from an unbalanced design, for example, different number of days at the sites. This is an important practical issue, since it may not always be possible to choose sites where there are several days of milling work. The difference from the balanced case is that the true values for the variance–covariance matrix for the variance estimates via REML are not the same as those from the sum of squares estimation method, and may be difficult to determine. However, as long as the standard error of each job mean accounts for a high percentage of the standard deviation of $E_1$, the same methodology seems reasonable. This percentage can be determined from the simulations. The simulations are used to adjust both for deviations from normality of $E_1$, and also for bias in the estimation of the variance–covariance matrix of the variance estimates. Simulations indicate that for mildly unbalanced data (relative to a balanced data set), the coverage is similar to that obtained for balanced data with the same variance parameters. For example, for a study with six sites and the number of days at each site given by $(1,1,1,2,1,3)$, simulation comparison of coverage is made relative to a study with six sites, two days at each site. Coverage is very similar for the same percentile for both balanced and unbalanced data, largely because the between site variance in the study is 3 times the within site variance.

In summary, the proposed method is an approximate method for obtaining the required upper confidence limit for the arithmetic mean. It involves selecting a range of true parameter values over which the required coverage is obtained. For every new data set, the simulations must be done to determine the nominal probability of the $t$ statistic to use to obtain the required confidence. Note that other approaches are available for determining upper confidence limits for the lognormal mean when there are multiple variance components. An example is the generalized confidence intervals method [Krishnamoorthy and Mathew 2002]. Much of that work has been developed for the one-way random effects model. However, some implementations of the method have been proposed for mixed models, though these implementations are not used here [Fonseca 2011]. Also a useful representation for mixed models is given in [Liao 2004].
SAS programs, for data on ln scale:

options pageno=1 nocenter;

*** Input Caterpillar Data ***;
data cat_data;
  input job $ site day rcs;
  if rcs > 0 then ln_rcs = log(rcs);
datalines;
GR 1 1 0.0409
GR 1 2 0.02891
GR 1 3 0.01633
GR 2 1 0.02116
GR 2 2 0.067
GR 2 3 0.0431
GR 3 1 0.0292
GR 3 2 0.0227
GR 3 3 0.0607
OP 1 1 0.04715
OP 1 2 0.04120
OP 1 3 0.02025
OP 2 1 0.02092
OP 2 2 0.04221
OP 2 3 0.06891
OP 3 1 0.0306
OP 3 2 0.08360
OP 3 3 0.11490
run;

*** Alternative: Input data from Best Practices Report Excel spreadsheet - 3 sites, 3 days per site ***;
data excel_data;
  input job $ site day rcs;
  if rcs > 0 then ln_rcs = log(rcs);
datalines;
GR 1 1 0.00507
GR 1 2 0.00349
GR 1 3 0.00250
GR 2 1 0.01200
GR 2 2 0.00307
GR 2 3 0.00495
GR 3 1 0.01010
GR 3 2 0.00575
GR 3 3 0.01480
OP 1 1 0.01320
OP 1 2 0.00605
OP 1 3 0.00972
OP 2 1 0.01520
OP 2 2 0.00674
OP 2 3 0.01070
OP 3 1 0.02380
OP 3 2 0.03000
OP 3 3 0.01690
run;

*** Two SAS programs are provided. Each program is a macro. The macro calls are at the end of the second program. ***;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
*** SAS Program 1: For data analysis on the natural log scale ***;
***************************************************************************;
%macro getEsts(dataset,pctl);

Macro getEsts will compute the UCL for the arithmetic means of two jobs evaluated simultaneously.

Parameters include:

DATASET is a SAS dataset with the following variables:

site    site number (1,2,3,...)

day     day number (1,2,3,...)

job     GR or OP

ln_rcs  respirable crystalline silica exposure (on log scale)

pctl    is a constant representing the probability of the t-distribution to use to give 97.5% UCLs (determined using simulation)

Using dataset: &dataset

Run MIXED procedure

proc mixed data=&dataset method=reml covtest asycov;
class job site day;
model ln_rcs = job / ddfm=satterth solution /*outp=my_pred outpm=my_predm residual influence*/;
random site day(site) / solution;
estimate 'op' intercept 1 job 0 1 / e;
estimate 'gr' intercept 1 job 1 0 / e;
ods output estimates=_estimates covparms=_covparms asycov=_asycov;
run;

Get estimates, standard errors, and error degrees of freedom for each job group

data _my_estimates(keep=est_op se_op df_op est_gr se_gr df_gr index);
retain                est_op se_op df_op est_gr se_gr df_gr;
set _estimates;
index=1;
if label='op' then do; est_op=estimate; se_op=stderr; df_op = df; end;
if label='gr' then do; est_gr=estimate; se_gr=stderr; df_gr = df; output; end;
run;

Get variance component estimates

data _my_covparms(keep=var_site var_day var_res index);
retain               var_site var_day var_res;
set _covparms;
index=1;
if CovParm='site' then do; var_site=estimate; end;
if CovParm='day(site)' then do; var_day=estimate; end;
if CovParm='Residual' then do; var_res=estimate; output; end;
run;

Get variances and covariances for variance components

data _my_asycov(keep=var_var_site var_var_day var_var_res cov_site_day cov_site_res cov_day_res index);
retain             var_var_site var_var_day var_var_res cov_site_day cov_site_res cov_day_res;
set _asycov;
index=1
if Row=1 then do; var_var_site=CovP1; cov_site_day=CovP2; cov_site_res=CovP3; end;
if Row=2 then do; var_var_day=CovP2; cov_day_res=CovP3; end;
if Row=3 then do; var_var_res=CovP3; output; end;
run;

Merge the three datasets (1 observation from each) together

data _my_calcs;
merge _my_estimates _my_covparms _my_asycov;
by index;
drop index;
run;

Calculate upper confidence limits using methods (a) and (b)
data _my_calcs;
set _my_calcs;
*** Estimate geometric means ***;
gm_op = exp(est_op);
gm_gr = exp(est_gr);
*** Estimate arithmetic means ***;
am_op = exp(est_op + 0.5*(var_site + var_day + var_res));
am_gr = exp(est_gr + 0.5*(var_site + var_day + var_res));
*** Estimate geometric standard deviations for variance components ***;
gsd_site = exp(sqrt(var_site));
gsd_day = exp(sqrt(var_day));
gsd_res = exp(sqrt(var_res));
*** Estimate variance of the natural log of the arithmetic mean ***;
var_ln_am_op = se_op**2 + 0.25*(var_var_site + var_var_day + var_var_res + 2*cov_site_day + 2*cov_site_res + 2*cov_day_res);
var_ln_am_gr = se_gr**2 + 0.25*(var_var_site + var_var_day + var_var_res + 2*cov_site_day + 2*cov_site_res + 2*cov_day_res);
*** UCL by first obtaining an UCL for ln(AM) [i.e., ln_am+t_val*sqrt(var(ln_am))], then exponentiating ***;
*** Note: here (a) uses pctl percentile of t distribution based on simulations ***;
ucl_a_op=am_op*exp(tinv(&pctl,df_op)*sqrt(var_ln_am_op));
ucl_a_gr=am_gr*exp(tinv(&pctl,df_gr)*sqrt(var_ln_am_gr));
*** Note: here (b) uses 97.5th percentile ***;
ucl_b_op=am_op*exp(tinv(0.975,df_op)*sqrt(var_ln_am_op));
ucl_b_gr=am_gr*exp(tinv(0.975,df_gr)*sqrt(var_ln_am_gr));
run;
proc print data=_my_calcs;
var df_gr df_op am_gr am_op gm_gr gm_op gsd_site gsd_day gsd_res ucl_a_gr ucl_a_op ucl_b_gr ucl_b_op;
title3 'Hand calculations of the AMs, GMs, GSDs, and UCLs';
title4 "Method a uses &pctl --- Method b uses 0.975";
run;
title1;
%mend;
*********************************************************************************
*** SAS Program 2: Simulation to determine empirical coverage of the UCL 97.5 ***;
***                for ln scale data                                          ***;
*********************************************************************************
%macro checkCoverage(scenarios,n_reps,pctl,seed1,seed2,seed3);
***************************************************************************;
*** Macro checkCoverage will check the coverage of confidence intervals ***;
*** for the arithmetic mean for two jobs evaluated simultaneously. ***;
***                                                                      ***;
*** Parameters include: ***;
***                                                                      ***;
*** SCENARIOS is a SAS dataset with the following variables: ***;
*** scenario scenario number (1,2,3...) ***;
*** gm_gr geometric mean for GROUNDSMAN ***;
*** gm_op geometric mean for OPERATOR ***;
*** var_s between-site variance (on natural log scale) ***;
*** if 0, will assume sd_s = 0.000001 ***;
*** var_d between-day variance (on natural log scale) ***;
*** var_r residual (within-day) variance (on natural log scale) ***;
*** n_reps is a constant representing the number of datasets to ***;
*** simulate per scenario ***;
***                                                                      ***;
*** pctl is a constant representing the percentile of the ***;
*** t-distribution to evaluate (in addition to 97.5) ***;
***                                                                      ***;
*** seed1 is the starting seed value for site effect ***;
*** seed2 is the starting seed value for day effect ***;
*** seed3 is the starting seed value for job effect ***;
***************************************************************************;

%macro checkCoverage(scenarios,n_reps,pctl,seed1,seed2,seed3);
***************************************************************************;
*** Macro checkCoverage will check the coverage of confidence intervals ***;
*** for the arithmetic mean for two jobs evaluated simultaneously. ***;
***                                                                      ***;
*** Parameters include: ***;
***                                                                      ***;
*** SCENARIOS is a SAS dataset with the following variables: ***;
*** scenario scenario number (1,2,3...) ***;
*** gm_gr geometric mean for GROUNDSMAN ***;
*** gm_op geometric mean for OPERATOR ***;
*** var_s between-site variance (on natural log scale) ***;
*** if 0, will assume sd_s = 0.000001 ***;
*** var_d between-day variance (on natural log scale) ***;
*** var_r residual (within-day) variance (on natural log scale) ***;
*** n_reps is a constant representing the number of datasets to ***;
*** simulate per scenario ***;
***                                                                      ***;
*** pctl is a constant representing the percentile of the ***;
*** t-distribution to evaluate (in addition to 97.5) ***;
***                                                                      ***;
*** seed1 is the starting seed value for site effect ***;
*** seed2 is the starting seed value for day effect ***;
*** seed3 is the starting seed value for job effect ***;
***************************************************************************;
data _my_scenarios;
set &scenarios;
beta0 = log(gm_op);
beta1 = log(gm_gr/gm_op);
sb = sqrt(var_b);
if sb = 0 then sb = 0.000001;
sd = sqrt(var_d);
sr = sqrt(var_r);
run;

proc print data=_my_scenarios;
run;

data _my_scenarios;
set _my_scenarios;
keep scenario beta0 beta1 sb sd sr;
run;

*** Simulate data ***;
data _my_sim_data;
set _my_scenarios;
 *** Assign starting seed values ***;
retain seed1 &seed1 seed2 &seed2 seed3 &seed3;
 *** Initialize random process ***;
call rannor(seed1,w1);
call rannor(seed2,w2);
call rannor(seed3,w3);
 *** Generate datasets ***;
do rep = 1 to &n_reps;
 *** Three sites ***;
do site = 1 to 3;
 *** Generate the random site effect ***;
call rannor(seed1,w1);
 *** Three days/site ***;
do day=1 to 3;
 *** Generate the random day effect ***;
call rannor (seed2,w2);
 *** Two jobs/day ***;
do job = 'GR','OP';
 *** Generate the random job effect ***;
call rannor (seed3,w3);
 *** Generate RCS on the natural log scale ***;
ln_rcs = beta0 + beta1*(job='GR') + sb*w1 + sd*w2 + sr*w3;
rcs = exp(ln_rcs);
output;
end;
end;
end;
run;

proc print data=_my_sim_data;
by scenario rep;
id scenario rep;
where rep in (1,2);
var beta0 beta1 sb sd sr site day job seed1 seed2 seed3 w1 w2 w3 ln_rcs rcs;
title 'Sample data';
run;

options nonotes; *** turn off notes;
ods exclude all;
proc mixed data=_my_sim_data method=reml asycov;
by scenario beta0 beta1 sb sd sr rep;
class site day job;
model ln_rcs = job / solution ddfm=satterth;
random site day(site) / solution;
estimate 'gr' intercept 1 job 1 0 / e;
estimate 'op' intercept 1 job 0 1 / e;
ods output estimates=_estimates covparms=_covparms asycov=_asycov;
run;
ods select all;
options notes; *** turn notes back on;

*** Get estimates, standard errors, and error degrees of freedom for each job group ***;
data _my_sim_estimates(keep=scenario beta0 beta1 sb sd sr rep est_op se_op df_op est_gr se_gr
df_gr);
retain est_op se_op df_op est_gr se_gr df_gr;
set _estimates;
by scenario beta0 beta1 sb sd sr rep;
if label='op' then do; est_op=estimate; se_op=stderr; df_op = df; end;
if label='gr' then do; est_gr=estimate; se_gr=stderr; df_gr = df; end;
if last.rep then do; output; end;
run;

*** Get variance component estimates ***;
data _my_sim_covparms(keep=scenario rep var_site var_day var_res);
retain var_site var_day var_res;
set _covparms;
by scenario rep;
if CovParm='site' then do; var_site=estimate; end;
if CovParm='day(site)' then do; var_day=estimate; end;
if CovParm='Residual' then do; var_res=estimate; end;
if last.rep then do; output; end;
run;

*** Get variances and covariances for variance components ***;
data _my_sim_asycov(keep=scenario rep var_var_site var_var_day var_var_res cov_site_day
cov_site_res cov_day_res);
retain var_var_site var_var_day var_var_res cov_site_day cov_site_res
cov_day_res;
set _asycov;
by scenario rep;
if Row=1 then do; var_var_site=CovP1; cov_site_day=CovP2; cov_site_res=CovP3; end;
if Row=2 then do; var_var_day=CovP2; cov_day_res=CovP3; end;
if Row=3 then do; var_var_res=CovP3; end;
if last.rep then do; output; end;
run;

*** Merge the three datasets (1 observation from each) together ***;
data _my_sim_calcs;
merge _my_sim_estimates _my_sim_covparms _my_sim_asycov;
by scenario rep;
run;

*** Calculate upper confidence limits using methods (a) and (b) ***;
data _my_sim_calcs;
set _my_sim_calcs;
 *** Estimate geometric means for simulated data ***;
gm_op = exp(est_op);
gm_gr = exp(est_gr);
 *** Estimate arithmetic means for simulated data ***;
am_op = exp(est_op + 0.5*(var_site + var_day + var_res));
am_gr = exp(est_gr + 0.5*(var_site + var_day + var_res));
lam_op=log(am_op);
lam_gr=log(am_gr);

*** Estimate geometric standard deviations for variance components for simulated data ***;
gsd_site = exp(sqrt(var_site));
gsd_day = exp(sqrt(var_day));
gsd_res = exp(sqrt(var_res));
 *** Estimate variance of the natural log of the arithmetic mean for simulated data ***;
\[ \text{var}_{\text{ln, am, op}} = \text{se}_{\text{op}}^2 + 0.25\left(\text{var}_{\text{var, site}} + \text{var}_{\text{var, day}} + \text{var}_{\text{var, res}} + 2\times\text{cov}_{\text{site, day}} + 2\times\text{cov}_{\text{site, res}} + 2\times\text{cov}_{\text{day, res}}\right) \]

\[ \text{var}_{\text{ln, am, gr}} = \text{se}_{\text{gr}}^2 + 0.25\left(\text{var}_{\text{var, site}} + \text{var}_{\text{var, day}} + \text{var}_{\text{var, res}} + 2\times\text{cov}_{\text{site, day}} + 2\times\text{cov}_{\text{site, res}} + 2\times\text{cov}_{\text{day, res}}\right) \]

*** Compute UCL from simulated data, using passed percentile ***;

\[ \text{ucl}_{\text{op, a}} = \text{am}_{\text{op}} \times \exp(tinv(\&pctl, \text{df}_{\text{op}}) \times \sqrt{\text{var}_{\text{ln, am, op}}}) \]

\[ \text{ucl}_{\text{gr, a}} = \text{am}_{\text{gr}} \times \exp(tinv(\&pctl, \text{df}_{\text{gr}}) \times \sqrt{\text{var}_{\text{ln, am, gr}}}) \]

*** Compute UCL from simulated data, using 97.5th percentile ***;

\[ \text{ucl}_{\text{op, b}} = \text{am}_{\text{op}} \times \exp(tinv(0.975, \text{df}_{\text{op}}) \times \sqrt{\text{var}_{\text{ln, am, op}}}) \]

\[ \text{ucl}_{\text{gr, b}} = \text{am}_{\text{gr}} \times \exp(tinv(0.975, \text{df}_{\text{gr}}) \times \sqrt{\text{var}_{\text{ln, am, gr}}}) \]

*** Compute true arithmetic means ***;

\[ \text{true}_{\text{am, op}} = \exp(\beta_0 + 0.5\times(\text{sb}^2 + \text{sd}^2 + \text{sr}^2)) \]

\[ \text{true}_{\text{am, gr}} = \exp(\beta_0 + \beta_1 + 0.5\times(\text{sb}^2 + \text{sd}^2 + \text{sr}^2)) \]

*** Determine if arithmetic mean is above UCL ***;

\[ \text{fraction}_{\text{op, a}} = \left(\text{true}_{\text{am, op}} < \text{ucl}_{\text{op, a}}\right) \]

\[ \text{fraction}_{\text{gr, a}} = \left(\text{true}_{\text{am, gr}} < \text{ucl}_{\text{gr, a}}\right) \]

\[ \text{fraction}_{\text{op, b}} = \left(\text{true}_{\text{am, op}} < \text{ucl}_{\text{op, b}}\right) \]

\[ \text{fraction}_{\text{gr, b}} = \left(\text{true}_{\text{am, gr}} < \text{ucl}_{\text{gr, b}}\right) \]

\[ \text{fraction}_{\text{op, a}, 05} = \left(\text{ucl}_{\text{op, a}} < 0.05\right) \]

\[ \text{fraction}_{\text{op, b}, 05} = \left(\text{ucl}_{\text{op, b}} < 0.05\right) \]

\[ \text{fraction}_{\text{gr, a}, 05} = \left(\text{ucl}_{\text{gr, a}} < 0.05\right) \]

\[ \text{fraction}_{\text{gr, b}, 05} = \left(\text{ucl}_{\text{gr, b}} < 0.05\right) \]

run;

proc means data=_my_sim_calcs n min max nmiss median mean var;
  class scenario;
  var true_am_op true_am_gr fraction_op_a fraction_gr_a fraction_op_b fraction_gr_b fraction_op_a_05 fraction_gr_a_05 fraction_op_b_05 fraction_gr_b_05;
  run;

*** Compute UCL from simulated data, using 97.5th percentile ***;

\[ \text{ucl}_{\text{op, b}} = \text{am}_{\text{op}} \times \exp(tinv(0.975, \text{df}_{\text{op}}) \times \sqrt{\text{var}_{\text{ln, am, op}}}) \]

\[ \text{ucl}_{\text{gr, b}} = \text{am}_{\text{gr}} \times \exp(tinv(0.975, \text{df}_{\text{gr}}) \times \sqrt{\text{var}_{\text{ln, am, gr}}}) \]

*** Compute true arithmetic means ***;

\[ \text{true}_{\text{am, op}} = \exp(\beta_0 + 0.5\times(\text{sb}^2 + \text{sd}^2 + \text{sr}^2)) \]

\[ \text{true}_{\text{am, gr}} = \exp(\beta_0 + \beta_1 + 0.5\times(\text{sb}^2 + \text{sd}^2 + \text{sr}^2)) \]

*** Determine if arithmetic mean is above UCL ***;

\[ \text{fraction}_{\text{op, a}} = \left(\text{true}_{\text{am, op}} < \text{ucl}_{\text{op, a}}\right) \]

\[ \text{fraction}_{\text{gr, a}} = \left(\text{true}_{\text{am, gr}} < \text{ucl}_{\text{gr, a}}\right) \]

\[ \text{fraction}_{\text{op, b}} = \left(\text{true}_{\text{am, op}} < \text{ucl}_{\text{op, b}}\right) \]

\[ \text{fraction}_{\text{gr, b}} = \left(\text{true}_{\text{am, gr}} < \text{ucl}_{\text{gr, b}}\right) \]

run;

%getEsts(cat_data,0.985);

data cat_scenarios;
  input scenario gm_gr gm_op var_b var_d var_r;
  datalines;
  1 0.033 0.045 0 0.17 0.12
  2 0.033 0.045 0 0.20 0.15
  3 0.033 0.045 0 0.30 0.20
  4 0.033 0.045 0 0.20 0.30
  5 0.033 0.045 0 0.10 0.10
  6 0.033 0.045 0 0.05 0.17 0.12
  7 0.033 0.045 0 0.05 0.20 0.15
  8 0.033 0.045 0 0.05 0.30 0.20
  9 0.033 0.045 0 0.05 0.20 0.30
  10 0.033 0.045 0 0.05 0.10 0.10
run;

%checkCoverage(cat_scenarios,5000,0.985,93022,4402,82003);
%getEsts(excel_data,0.995);
data excel_scenarios;
    input scenario gm_gr gm_op var_b var_d var_r;   *** not much thought given to these, for
data-only...;
datalines;
1 0.013 0.0058 0.033 0.19 0.077 0.12
2 0.013 0.0058 0.033 0.30 0.100 0.12
3 0.013 0.0058 0.033 0.30 0.200 0.12
4 0.013 0.0058 0.033 0.10 0.100 0.12
run;
%checkCoverage(excel_scenarios,50,0.995,0,0,0);
/**/

References:
Fonseca, M., Mathew, T., Mexia, J., and Zmylony, R. [2011]. AIP Conference

Krishnamoorthy K, and Mathew T [2002]. Assessing occupational exposure via the
one-way random effects model with balanced data. Journal of Agricultural,

for general balanced mixed models and unbalanced one-way random

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., and Schabenberger,

Mandel M [2013]. Simulation-based confidence intervals for functions with
complicated derivatives. The American Statistician. 67(2), pp. 76.

Inc.

York: Wiley, pp. 63-64.
Delivering on the Nation’s promise: Safety and health at work for all people through research and prevention.

To receive NIOSH documents or other information about occupational safety and health topics, contact NIOSH at

1-800-CDC-INFO (1-800-232-4636)
TTY: 1-888-232-6348
E-mail: cdcinfo@cdc.gov

or visit the NIOSH Web site at www.cdc.gov/niosh

For a monthly update on news at NIOSH, subscribe to NIOSH eNews by visiting www.cdc.gov/niosh/eNews

SAFER ● HEALTHIER ● PEOPLE