2012

# Limitations associated with the pooling of historical data sets: an analysis of dust monitoring data collected at a typical remote Australian mine 2004 - 2008

Jacques Oosthuizen
*Edith Cowan University*

# [¶28-221] Limitations associated with the pooling of historical data sets: an analysis of dust monitoring data collected at a typical remote Australian mine 2004 — 2008

Click to open document in a browser

Last reviewed: 18 October 2012

***J Oosthuizen***

Jacques Oosthuizen (PhD), Edith Cowan University, School of Exercise, Biomedical and Health Sciences, Perth, Australia.

Address for correspondence: Dr Jacques Oosthuizen, Edith Cowan University, School of Exercise Biomedical and Health Sciences, 270 Joondalup Drive, Joondalup, WA 6027. j.oosthuizen@ecu.edu.au

**Keywords: dust sampling, aggregate data, mining**.

**Abstract**

**Dust monitoring data collected over a period of 5 years at a remote Australian mine site were pooled in order to create a large data set for analysis. Results adjusted for a 12-hour shift were compared to exposure standards and 90% of the respirable and 97% of the inhalable dust samples were found to be within the exposure standards.**

**Median values appear to better describe exposures than the mean which is influenced by outliers. There was a declining trend in respirable dust results over time. However, inhalable dust levels remained fairly constant over the sampling period. Inhalable dust levels were higher during the winter, yet respirable dust was elevated in the spring.**

**Regulators need to work with practitioners in order to design dust monitoring protocols that address sampling limitations and assist mining operations to fulfil their obligations in terms of occupational hygiene monitoring and protection of their workers.**

**Introduction**

Exposure surveillance in the workplace should be conducted in terms of systematic, ongoing pre-established protocols that are determined through the evaluation of current and historical occupational hygiene data. Exposure monitoring protocols should detail clearly what needs to be sampled and how samples should be collected. Sample collection handling and analysis needs to be conducted according to standardised methods that will ensure uniformity, thus enabling the comparison of data over time, particularly in situations where samples are collected by more than one person.[1,2]

Furthermore exposure standards utilised need to be valid and should provide adequate protection. In situations where people are employed on a 12-hour shift for example, or in situations where they may be exposed to multiple agents, exposure standards may require adjustment. Details such as sample type (personal or area), sample duration (run time) and specification of the analytical methods used are also frequently lacking in historical data sets. Additional variations that may introduce bias include,

- the use of different sampling instruments
- use of incorrect filters with differing pore size
- inaccurate instrument calibration
- inappropriate (non-random) selection of sampling days, shifts and people

- length of time and time of day sampled. Exposures can be overestimated if samples are collected for a short time period during high production, or underestimated if collected for a limited period that includes a lunch break
- length of shift and appropriateness of the exposures standard or adjusted standard
- different sampling strategies, and different analytical methods.[3, 4, 5]

There are a variety of different devices that have been used for respirable dust monitoring over a number of years; these include impingers, konimeters and gravimetric methods. Although gravimetric sampling has been conducted for many decades, there were significant variations in methods used until these were standardised internationally in 2000. Previously open-faced filters were frequently used, thus comparison of respirable dust results over time can be problematic. An additional source of concern in epidemiological exposure assessment is the fact that dust surveys are normally conducted for occupational hygiene/ compliance testing rather than for epidemiological exposure assessment. Worst case scenario sampling is often conducted that may not provide a true reflection of the occupational exposures experienced within particular cohorts.[4]

In 2004 the Australian standard prescribing respirable dust sampling was updated to reflect trends internationally. The standard prescribed a gravimetric method utilising a cyclone elutriator. The sample flow rate was increased from 1.9 to 2.2 litres per minute for BCIRA and SIMPEDS cyclones and the Aluminium cyclone flow rate was set at 2.5 L/min. Samples are to be collected over a minimum period of four hours, thus comparisons of results collected prior to and after 2004 need to take this into account. In 2009 a further revision of the standard was done; however, sample methodology remained unchanged.[5]

### Background and rationale for the review

It is usually difficult in an operational context, to gather the large numbers of measurements necessary to produce a high level of confidence in exposure data, particularly when there is significant variability in the results, or the mean exposure level is close to the exposure standard[6] .

Normal practice at the mine site where the data considered in this paper was obtained was to collect approximately 6–10 samples for every similar exposure group (SEG) and decisions regarding exposure assessment were thus based upon small sample sizes. The Occupational Safety and Health manager at the site in question was concerned that exposures in various areas of the operation had not been adequately quantified in the past and the objective of this review was to increase confidence in the data by compiling a large aggregate data set, thus enabling an analysis of trends over time and the development of a five-year sampling strategy.

This paper explores issues related to the aggregation of large data sets and in particular the limitations associated with subsequent analysis. Although a more detailed analysis for specific contaminants and similar exposure groups (SEGs) was performed, permission to publish these data was not granted. This review is restricted to a general discussion of aggregated respirable and inhalable dust results.

### Methods

All samples of airborne contaminants that had been collected at the mine site during the period 2004–2009 were pooled to create an MS Excel spreadsheet[7] consisting of 2,226 undifferentiated gravimetric dust results. Dust samples were collected by a variety of different people using different instruments and in some instances details related to date, shift, sample run time, filter type and exact location were not recorded. Samples were shipped to a National Association of Testing Authorities (NATA) approved laboratory for analysis.

The data set was filtered to differentiate between inhalable (792) and respirable samples (1,434), while 238 results were discarded as they were not identified as being either respirable or inhalable. A number of results had notes attached indicating problems that were encountered during sampling, such as damaged filters or pumps cutting out. Furthermore, some samples were marked with one or more asterisks (*) with no further explanation regarding the meaning of the *. Results that were found to be below the level of detection were indicated with a < symbol. These results were manipulated by multiplying the level of detection by 0.7 in order to obtain a value that could be used for statistical analysis. Since the standard deviation for both respirable and inhalable dust was less than 3, this factor (0.7), was deemed appropriate.[6]

The two data sets were subsequently differentiated further in order to determine if there were seasonal variations in dust levels. Samples collected during May to August were classified as winter, September to November (spring), December to February (summer) and March & April were classified as autumn.

In an attempt to evaluate the effectiveness of dust suppression strategies implemented over time, the data sets were also differentiated according to the year samples were collected and these data were analysed to identify if there was variability over the six-year sampling period.

Dust sampling results were compared to the exposure standards published by Safe Work Australia.[8] These exposure standards were adjusted for a 12-hour shift to 5mg/m$^3$ (inhalable) and 1.5 mg/m$^3$ (respirable), according to the Brief and Scala method.[9]

**Results**

**Respirable**

Respirable samples (n = 1434) had a mean value of 1.08 mg/m$^3$ and 10% of samples exceeded the exposure standard of 1.5 mg/m$^3$. However, there were a few outliers in the data set that were not adequately explained other than being annotated with an (*).

The maximum value of 250 mg/m$^3$ is unlikely to be accurate as are 11 values greater than 30 mg/m$^3$, particularly since 97% of the results were below 5 mg/m$^3$. Due to the vast discrepancy in the upper end of the range which is likely due to sampling errors, values above 30 mg/m$^3$ were deleted to create an adjusted data set. The adjusted mean value was reduced from 1.08 to 0.66 mg/m$^3$ as demonstrated in Table 1. It is interesting to note that the median and the mode remain unaffected at 0.2 and 0.07 mg/m$^3$ respectively and the median is in fact close to the adjusted mean value of 0.66 mg/m$^3$. The mode appears to underestimate exposures; however, this result was influenced by the large proportion of samples found to be below the level of detection (31%), thus rendering the mode invalid as a measure of central tendency in this case.

**Inhalable**

Fifty-three different dust levels were measured in a data set of 792 samples ranging from 0.07 mg.m$^{-3}$ to 111.4 mg.m$^{-3}$. The majority of the results (79%) were below 1 mg.m$^{-3}$ and 96% of samples were below the exposure standard of 5 mg/m$^3$. However, there were some results that were extremely high and not likely to be correct. For example there were eight results between 30 and 33 mg/m$^3$ and two above 110 mg/m$^3$. When these 10 outliers are removed from the data set the adjusted mean is reduced from 1.31 to 0.84, increasing compliance to 97%. The adjusted mean is also much closer to the median value of 0.2, while the mode appears to underestimate the true exposure at 0.07; again this is related to the large number of samples below the level of detection (32%) (see Table 1).

**Table 1: Aggregate dust sampling results**

| Sample | Min mg/m$^3$ | Max mg/m$^3$ | Mean | SD | 95% CI | Median mg/m$^3$ | Mode mg/m$^3$ | N | % OES |
|---|---|---|---|---|---|---|---|---|---|
| Respi-rable | 0.04 | 250 | 1.08 | 7.37 | 0.95–1.21 | 0.2 | 0.07 | 1434 | 10 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.04 | 23 | 0.66 | 1.78 | 0.63–0.69 | 0.2 | 0.07 | 1422 | 6 |
| Inhale-able | 0.07 | 111.4 | 1.31 | 6.42 | 1.16–1.46 | 0.2 | 0.07 | 792 | 4 |
| | 0.07 | 18 | 0.84 | 2.17 | 0.79–0.89 | 0.2 | 0.07 | 782 | 3 |

% OES – percentage over exposure standard

**Seasonal variation**

Data for both inhalable and respirable dust were further analysed in order to determine if there was any detectable seasonal variation in dust levels (Table 2). Sixty-eight respirable samples were excluded from the analysis as dates of collection were not specified. These samples had a mean value of 0.45 (SD 1.24, CI 0.79–1.69), a maximum of 9.9 mg/m$^3$ and 3% of these samples exceeded the exposure standard. Respirable dust levels measured during spring, the driest month (n = 201) were significantly higher (p = 0.005) than for the other seasons combined with a mean value of 1.15 mg/m$^3$ (SD 3.03, CI 1.29–1.01).

## Table 2: Seasonal variation in respirable dust levels

| Season | Min mg/m$^3$ | Max mg/m$^3$ | Mean mg/m$^3$ | SD | 95% CI | Median mg/m$^3$ | Mode mg/m$^3$ | N | % OEL |
|---|---|---|---|---|---|---|---|---|---|
| Autumn | 0.07 | 18 | 0.74 | 2.15 | 0.66–0.82 | 0.2 | 0.07 | 354 | 8 |
| Winter | 0.04 | 6.1 | 0.63 | 1.01 | 0.58–0.68 | 0.2 | 0.07 | 199 | 13 |
| Spring | 0.07 | 23 | 1.15 | 3.03 | 1.29–1.01 | 0.07 | 0.07 | 201 | 17 |
| Summer | 0.07 | 14 | 0.49 | 1.01 | 0.46–0.52 | 0.2 | 0.04 | 600 | 7 |

% OES – percentage over exposure standard

Results for inhalable dust, stratified by season, are presented in Table 3; 49 results were excluded as the date of sampling was not recorded. The discarded samples had a mean value of 0.29 mg/m$^3$ (SD 0.47, CI 0.24–0.34), and a maximum value of 1.3 mg/m$^3$.

## Table 3: Seasonal variation in inhalable dust levels

| Season | Min mg/m$^3$ | Max mg/m$^3$ | Mean mg/m$^3$ | SD | 95% CI | Median mg/m$^3$ | Mode mg/m$^3$ | N | % OEL |
|---|---|---|---|---|---|---|---|---|---|
| Aut | 0.07 | 0.2 | 0.09 | 0.04 | 0.08–0.1 | 0.07 | 0.07 | 261 | 0 |
| Win | 1.4 | 16 | 3.74 | 3.48 | 3.49–3.99 | 2.2 | 1.4 | 90 | 18 |
| Spr | 0.2 | 0.2 | 0.2 | | | 0.2 | 0.2 | 20 | 0 |
| Sum | 0.2 | 1.2 | 0.48 | 0.27 | 0.47–0.49 | 0.4 | 0.2 | 352 | 0 |

% OES – percentage over exposure standard

**Annual trends**

Respirable and inhalable dust data were analysed separately for each year in order to determine if there were any detectable variations over time. A total of 49 results were excluded as the year of sampling was not specified.

Respirable dust results over the six-year period appear to show a declining trend. The mean results of 2004 and 2005 are significantly higher than in subsequent years (p = 0.03). However, in 2008, levels were again elevated to 0.98 mg/m$^3$ (Figure 1).

**Figure 1: Mean respirable dust levels for 2004–2009**

MS 1248 Fig1

Mean inhalable dust levels remain fairly constant for the period 2005 to 2009 (Figure 2)

**Figure 2: Mean annual inhalable dust results (2005–2009 — no data for 2004)**

MS 1248 Fig2

### Discussion

Ideally, occupational hygiene sampling strategies should be based upon random sampling techniques. True random sampling includes more than the mere selection of workers to be included in a survey. Technically the sampling strategy should also include random selection of days of the year and work shift sampled. Typical remote mining operations normally have dynamic workplaces causing exposures to fluctuate throughout a normal work day or at certain stages of production and maintenance cycles. One of the ways to overcome the potential errors or uncertainty surrounding exposure data is to take many more samples. Due to cost and time constraints this is usually not a viable option in the short-term. The next best alternative is to aggregate data that has been collected over a long period of time and to do an analysis of the larger data set.[1]

Historical data is frequently collected by different people over time, leading to uncertainty associated with possible variations in the sampling methods, instruments and analytical techniques used, as well as in data interpretation. One of the most significant sources of variation is the treatment of samples below the level of detection, particularly since a large portion of the results fall into this category. It is therefore important that industries should adopt standard sampling and reporting protocols that should be subjected to external peer review in order to ensure scientific rigour. Samples collected in an industrial setting over a long period of time can contribute significantly to knowledge related to worker exposures and can be linked to health surveillance data thus allowing for occupational epidemiological evaluations.

From a regulatory perspective this study site appears to be reasonably compliant. However, there are a number of limitations typical of historical data analysis that make interpretation of the results problematic. Many different people had been responsible for the collection of dust samples over the six-year period, thus there was not much uniformity in the way results were recorded and presented. Many of the data sets were incomplete and lacked information related to date of sampling, specific area where samples were collected, run time, flow rates, filter types, analytical procedures, problems encountered during sampling, sample method (respirable/inhalable), equipment used and calibration details. It is important to note that the prescribed flow rate for respirable dust monitoring changed during this six-year period and the point at which the new flow rate was adopted is not clear.[1]

### Aggregate data

More than 2,000 dust samples were deemed acceptable for inclusion in a historical analysis (n = 2,204). Respirable dust (n = 1,422) with a mean value of 0.66 mg/m$^3$, was compared to the 12-hour adjusted exposure standard of 1.5 mg/m$^3$. Due to a number of reasonable high results the range of values (0.04–23) was wide. The standard deviation (SD) was 1.78 and 95% confidence interval (CI) 0.63–0.69. Overall compliance in terms of the exposure standard was reasonably good (94%).

The inhalable dust (n = 782) samples had a mean value of 0.84mg/m$^3$ which is well below the 12-hour shift adjusted exposure standard of 5 mg/m$^3$. However, the range was wide due to a small number of high values (range from 0.07 to 18). The SD was 2.17 and the CI (0.79–0.89) with 97% of samples being lower than the exposure standard.

On the whole, dust monitoring results at the sample location appear to be acceptable. It is probable that the majority of the exceedences were associated with specific tasks during maintenance or shutdown. It was reported anecdotally that targeted “worst case scenario” sampling during specific maintenance “shutdown” procedures had been conducted at this operation over the years and employees were required to wear dust masks during shutdowns; these results were incorporated into the aggregate data set.

### Seasonal variation

The mine site is located in an arid region with naturally occurring (red dirt) dust during dry windy periods; this situation is exacerbated by the mining activities. Long-term rainfall data sourced from the Australian Bureau of Meteorology.[10] weather station closest to the mine are summarised in Table 4.

## Table 4: Summary of average seasonal rainfall in the region for the last 30 years

| Autumn | Winter | Spring | Summer |
|---|---|---|---|
| 43 mm | 62 mm | 40 mm | 127 mm |

Respirable dust levels measured during spring, the driest month (n = 201) were significantly higher (p = 0.005) than for the other seasons combined with a mean value of 1.15 mg/m$^3$ (SD 3.03, CI 1.29–1.01). The greatest number of exceedences of the exposure standard also occurred during the spring (17%) and it is possible that dry windy conditions contributed to this finding.

Inhalable dust (n = 90) was significantly elevated during the winter (p = 0.0006) with a mean value of 3.74 mg/m$^3$ (SD 3.48, CI 3.49–3.99) and 18% exceedences of the exposure standard.

### Annual trends

Respirable dust results over the six-year period appear to show a declining trend. The mean results of 2004 and 2005 are significantly higher than in subsequent years (p = 0.03). However, in 2008, levels were again elevated to a mean of 0.98 mg/m$^3$.

Mean inhalable dust levels remain fairly constant for the period 2005 to 2009. Exceedences of the exposures standard were 0% (2005 and 2009), 3% (2006 and 2008) and 6% in 2007.

Across all the sub-groups analysed both the median and the mode were lower than the mean values (Table 5). In this data set the mean and median values are closely correlated (r = 0.88), while correlation between the mean and the mode is positive but not at strong (r = 0.79). The median and mode are more closely correlated at (r = 0.89).

The median may be a useful descriptor of average exposures in the workplace over time as it is not influenced by the outliers on the upper end of the data set and it is a frequently used measure of central tendency in respirable dust surveys.[11, 12, 13] Mode may underestimate exposures as most results were at the lower end of the scale, or below the level of detection, thus these were the most commonly occurring values.

## Table 5: Comparison between mean median and mode as a descriptor of exposure

| Data set | Mean | Median | Mode |
|---|---|---|---|
| Aggregate Respirable | 0.66 | 0.2 | 0.07 |
| Aggregate Inhalable | 0.84 | 0.2 | 0.07 |
| Seasons | 0.74 | 0.2 | 0.07 |
| Respirable | 0.63 | 0.2 | 0.07 |
| | 1.15 | 0.07 | 0.07 |
| | 0.49 | 0.2 | 0.04 |
| Seasons | 0.09 | 0.07 | 0.07 |
| Inhalable | 3.74 | 2.2 | 1.4 |
| | 0.2 | 0.2 | 0.2 |
| | 0.48 | 0.4 | 0.2 |
| Annual | 1.15 | 0.1 | 0.5 |
| Respirable | 1.27 | 0.4 | 0.07 |

| | | | |
|---|---|---|---|
| | 0.66 | 0.1 | 0.07 |
| | 0.32 | 0.1 | 0.07 |
| | 0.98 | 0.2 | 0.07 |
| | 0.4 | 0.2 | 0.07 |
| Annual | 1.2 | 0.6 | 0.07 |
| Inhalable | 0.73 | 0.2 | 0.07 |
| | 1.12 | 0.3 | 0.2 |
| | 0.89 | 0.5 | 0.2 |
| | 1.05 | 0.9 | 0.9 |

**Conclusion**

The impact of climatic conditions (dry windy conditions and dust storms) on occupational dust samples is not known. However, it is likely that respirable dust levels are influenced by weather conditions. Mining activities will exacerbate the amount of windblown dust and it may be necessary to survey dust levels prior to the commissioning of new mine sites in order to establish naturally occurring baseline data. This would be particularly significant where communities reside in close proximity to proposed mining developments in order to quantify the relative contribution of the mining operation to overall dust exposure in the area. In any event, from an occupational health and safety perspective workers need to be protected adequately from exposures at their place of employment irrespective of the source.

The analysis of annual aggregated dust levels over time can provide a useful measure of the efficiency of dust suppression controls and as such it could be an important tool for use by regulators and practitioners. The problem with this particular data set is that sampling was conducted at different times of the year for each year, and particularly during shutdowns, thus making direct comparisons of annual data problematic. If this approach is to be effective the annual dust monitoring protocols should be repeated each year to be more closely related in terms of time of the year sampled, with rainy days and dust storm conditions clearly highlighted for separate analysis and comparison.

A significant limitation of long-term monitoring in remote locations appears to be related to the high turnover of staff and lack of continuity in implementing a long-term scientifically valid sampling protocol. There is generally a lack of clearly described sampling methods, areas and techniques.

Regulators need to work with practitioners in order to assist mining operations to design sampling protocols that will:

- protect workers
- address limitations highlighted in this review
- contribute to exposure profiling data of occupational cohorts for epidemiological purposes, and
- fulfil statutory obligations.

Good record-keeping and communication during handover of duties is an essential requirement for the success of a valid sampling program that can be used to quantify exposures from an epidemiological perspective.

**References**

**Footnotes**

1 Oosthuizen, J. Introduction to occupational hygiene monitoring for technicians. Lulu Enterprises Inc, USA 2006, pp 8–12.

2 Vincent JH, Werner MA. Critical evaluation of historical occupational aerosol exposure records: Applications to Nickel and Lead. Annals of Occupational Hygiene 2003; 47(1): 49–59.

3 Di Nardi SR, ed. The occupational environment: Its evaluation, control and management 2nd edn. Fairfax VA: AIHA Press, 2003, p 1246.

4 Dahmann D, Taeger D, Kappler M, Buchte S, Morfeld TB, Pesh, B. Assessment of exposure in epidemiological studies: the example of silica dust. Journal of exposure science and environmental epidemiology 2008;18:452–461.

5 AS 2985–2009, Workplace atmospheres — Method for sampling and gravimetric determination of respirable dust. Sydney, Standards Australia 2009.

6 Mulhausen JR, Damiano J. A strategy for assessing and managing occupational exposures. Fairfax VA: AIHA Press, 1998, p 145.

7 Microsoft Excel 2007. Available at http://www.microsoft.com/en/us/default.aspx. Accessed 2010

8 Safe Work Australia (2009) Canberra: Available at http://www.safeworkaustralia.gov.au/swa/. Accessed 2010

9 Brief RS, Scala RA. Occupational Exposure Limits For Novel Work Schedules. American Industrial Hygiene Association Journal 1975;36(6):467–469.

10 Australian Bureau of Meterology (2010) Canberra. Available at: http://www.bom.gov.au/climate/averages/tables/cw_012046.shtml. Accessed 2010

11 Spies A, Rees D, Fourie AM, Wilson KS et al. Polybrominated Diphenyl Ethers in Indoor Dust in Ottawa, Canada: Implications for Sources and Exposure. International Journal of Occupational and Environmental Health 2008;14(3):225–321.

12 Wilford BH, Shoeib H, Harner T, Zhu J, Jones KC. Personal Exposure to Polybrominated Diphenyl Ethers (PBDEs) in Residential Indoor Air. Environmental. Science and Technology 2005;39(18):7027–7035.

13 Elms J, Robinson E, Rahman S, Garrod A. Exposure to flour dust in UK bakeries: Current use of control measures. Annals of Occupational Hygiene 2005;49(1):85–91.