Method-induced Misclassification for a Respirable Dust Sampled Using ISO/ACGIH/CEN Criteria

D. L. JOHNSON* and N. A. ESMEN

Aerosols Research Laboratory, University of Oklahoma Health Sciences Centre, PO Box 26901, Oklahoma City, OK 73190, USA

Received 22 April 2002; in final form 23 May 2003

The single cut measurement of exposure to respirable dust is the accepted method of exposure classification in occupational hygiene. We previously showed that actual pulmonary tissue dose may be substantially different from the dose expected, or the indicated dose, based on measurements using current single cut methodologies. We now examine exposure misclassification of workers based on any single cut respirable dust measurement using the internationally accepted ISO/ACGIH/CEN single cut respirable dust measurement criteria. Hypothetical aerosols with 12 generalized size distributions typical of the method of aerosol generation (condensation, aged condensation, mechanical low energy, mechanical high energy and mixtures thereof) were assumed. Using previously reported models for sampler penetration and pulmonary deposition, Monte Carlo simulations of actual mass dose to pulmonary tissues in comparison to the dose estimate from an ideal respirable aerosol sampler were carried out. Measurement-based indicated doses were used to classify exposures into five exposure categories and these classifications were compared with the ‘true’ classifications from the dose-based exposure estimates. Misclassification rates were generally severe and were greatest for aerosols with mass median aerodynamic diameter (MMAD) < 1 \(\mu\)m (\(\sim 100\%\)) and MMAD 5–15 \(\mu\)m (65–95\%). Misclassification rates were moderate (<20\%) only for extremely coarse aerosols of MMAD > 15 \(\mu\)m. Misclassification rates for oral and nasal breathing at 750 and 1500 ml tidal volume and 15 breaths/min were similar for each aerosol examined.

Keywords: size-selective sampling; respirable; error; misclassification

INTRODUCTION

The measurement of mass concentration of particles by a pre-classifier that mimics the penetration of dust to the non-ciliated regions of the respiratory track as the metric for exposure to respirable dust is the accepted method used for one type of exposure classification in occupational hygiene. Generally, exposures to pneumoconiotic dusts are characterized by this metric. The underlying supposition is that this exposure measure relates in some consistent way to dose. We previously showed that actual pulmonary tissue dose might be substantially different from the dose expected based on measurements using current single cut methodologies (Esmen and Johnson, 2002), in agreement with findings by others. For example, Vincent and Mark (1984) identified biases in estimates of alveolar coal dust deposition for a given aerosol under different breathing conditions when using the BMRC respirable dust convention. Hewett (1991) also demonstrated this significant particle size distribution-dependent bias for the respirable as well as inhalable and thoracic size fractions and cautioned that the associated variability introduced into dose–response analyses could lessen the chance of observing a statistically significant dose–response relationship. Additionally, McCawley (1993, 1999) has discussed in detail the potential for bias in alveolar deposition estimates for respirable aerosols with significant sub-micrometre particle fractions. These findings reflect the physical reality that not all particles that are inhaled are deposited, but rather may remain suspended and be subsequently exhaled. The mass fraction of inhaled aerosol that deposits in the pulmonary region is influenced by a number of factors, including breathing frequency, breath volume and mouth versus nose breathing (all

*Author to whom correspondence should be addressed. Tel: +1-405-271-2070 ext. 46776; fax: +1-405-271-1971; e-mail: David-Johnson@ouhsc.edu
of which are related to work rate), particle size distribution and the individual’s respiratory tract morphology. These influences are discussed in some depth in our previous work and will not be repeated here. In this work we examine exposure misclassification of workers based on the accepted and widely used respirable dust measurement metric.

In an historical context, the use of methods based on the penetration of dust as the exposure metric were proposed as an improvement upon other dust exposure indices such as impinger- and konimetre-based measurements (Walton and Vincent, 1998). These older measurements could not be readily or at all related to the particulate matter dose received in the air exchange regions of the human lung. Consequently, starting from an explicitly expressed concern by Hatch and Hemeon (1948), through the development of sampler theory (Walton, 1954) and static (Wright, 1954) and personal sampling instrumentation (Lippmann and Harris, 1962), particle penetration-based exposure measurement became the accepted exposure metric. Indubitably, these penetration-based measurement techniques represented a great improvement in the epidemiologic predictability of pneumoconiotic diseases and led to a significant improvement in exposure–response relationships. It must be understood that the developers of these sampling conventions appreciated the limitations of penetration-based size selective sampling criteria and recognized that they represented a necessary trade-off between actual aerosol inhalation and deposition behaviour and the limitations of available air sampling technology in simulating that behaviour (Walton and Vincent, 1998). However, with major improvements in the levels of dustiness over the intervening half-century, the possibility of the inadequacy of the use of a penetration-based metric in industrial epidemiology, in particular, must be considered. This is based on two somewhat related observations. If the exposures measured were such that the highest exposures that are well within the capabilities of the analytical techniques indicate the frank presence of disease, then these exposures would be dominant in the determination of an exposure–response relationship with respect to exposures that are relatively low due to absolute particle concentration. On the other hand, if all exposures are low, the gradation between exposure classes becomes limited by the capabilities of the analytical techniques and the ‘high’ and ‘low’ exposures are at levels that would have been considered to be the same exposure level in the past. In other words if high exposure level is of the order of 100 and the low exposure level is of the order of 1 in a previous study and exposures are at a high level of the order of 1 and at a low level of the order of 0.1 in a later study, all exposures of the latter study would be more or less in the same exposure class of the former study. Consequently, the gradation in exposure would require much more precise and effect-related measurements. Second is a somewhat disconcerting observation that in the traditionally dusty trades, dramatic reductions in dust exposure levels notwithstanding, there is still a residual presence of pneumoconiotic disease (NIOSH, 1995). Although there might be other explanations, unless otherwise shown the potential explanation, that an inherent misclassification of worker exposures is the confounding factor, cannot be readily dismissed.

In a recent paper, we examined the magnitude of the inhaled aerosol size distribution-based discrepancy between a penetration-based exposure metric and a deposition-based exposure metric. Even though the results of this study showed considerable disagreement between the exposure levels computed by each method, this incongruity would be epidemiologically acceptable if exposure classification by one method could be homeomorphically mapped onto the exposure classification by the other. In other words, an exposure classified by one of the methods would have the same rank if it were to be classified by the other. From the results reported in the previous paper, this requirement could not be readily confirmed or disputed. Therefore, in this work an investigation of potential misclassification rates for respirable dust exposures expressed by the internationally accepted single cut respirable dust exposure measurement method is reported.

**METHODS**

Hypothetical aerosols with 12 generalized ranges of size distributions typical of the method of aerosol generation (condensation, aged condensation, mechanical low energy, mechanical high energy and mixtures thereof) were assumed. These aerosols were log-normally distributed with mass median aerodynamic diameters (MMAD) of 0.1–35 µm and geometric standard deviations (GSD) of 1.8–3.5. Using previously reported models for sampler penetration and pulmonary deposition (Heyder et al., 1985, 1986; Rudolf et al., 1986, 1988), Monte Carlo simulations of actual mass dose to pulmonary tissues versus the dose estimate from an ideal respirable sampler were carried out using software developed specifically for the purpose. Each aerosol was characterized with a typical range of MMAD and GSD (Table 1). One-hundred combinations of MMAD and GSD were randomly selected from within the MMAD and GSD ranges specified for each aerosol and Romberg–Richardson numerical integrations were performed as previously described. The efficiency of particle mass deposition in pulmonary tissues as predicted from the empirical equations and the efficiency of particle mass penetration through an ideal respirable aerosol sampler as predicted by the ISO/ACGIH/CEN sampler performance criteria were
Method induced misclassification

Simulations were performed for both oral and nasal breathing and 750 and 1500 ml tidal volumes, at a breathing rate of 15 breaths/min. Indicated doses, represented by the mass penetrations $P_i$ through the respirable sampler, were used to classify exposures into categories arbitrarily selected as very low, low, moderate, high and very high corresponding to mass penetration efficiencies of <0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8 and ≥0.8.

'True' doses ($D_i$), represented by the mass deposition predicted from the empirical equations, were also ranked into these five categories. Deposition and penetration efficiencies were used rather than actual masses in order to remove aerosol concentration as a variable. The data were then examined to determine whether the penetration-based exposure classification corresponded to the 'true' classification that would have resulted from using deposition efficiency estimates. A trial in which the penetration-based classification did not match the deposition-based classification was judged to be misclassified. The absolute misclassification rate for each exposure category was then the fraction of misclassified exposures present in each category.

It was recognized that while exposures could be badly misclassified in terms of the actual doses received, as micrograms of particulate matter per kilogram of tissue perhaps, relative agreement between penetration- and deposition-based classifications in terms of membership of a 'level' could still be useful since only a scaling factor would distinguish the dose estimates. Therefore, a second evaluation was performed to determine the relative rate of misclassification between category assignments based on indicated exposures and assignments based on true doses. In this case the measured exposures were classified as before, but the 'true' doses were assigned to five equal width categories spanning the more restricted range of deposition values observed from the numerical simulations. A five level classification scheme was selected as a conveniently small but sufficiently general classification scheme. A misclassification metric $\Xi$, based on both the number of misclassified attributions and the level of misclassification for each aerosol type, may be calculated as (see Appendix):

$$\Xi = \frac{\sum_{i=1}^{m} |D_i - P_i|}{m(N-1)}$$  \hspace{1cm} (1)
For an aerosol of a given MMAD and GSD, dose $D_i$ and penetration $P_i$, five classes ($N = 5$), which ranged between 1 (very low) and 5 (very high), were the selected based on deposition and penetration, respectively. Monte Carlo simulations ($m = 100$) for each aerosol type were performed. The mean value of the deviation for misclassified exposures was also calculated for each aerosol type as:

$$\delta = \frac{1}{k} \sum_{i=1}^{m} |D_i - P_i|$$  \hspace{1cm} (2)

with $k$ as the number of misclassified Monte Carlo trials for the aerosol type.

RESULTS

Calculated penetration and deposition efficiencies for each of the 1200 simulations for mouth breathing with 750 ml tidal volume are shown in Fig. 1. The plot reflects the form of the ISO/ACGIH/CEN (penetration) criterion curve and the Heyder et al. equation (deposition) curve, however, the efficiency values in this case are for polydisperse rather than monodisperse aerosols. The variation in efficiency values about a given MMAD is due to the influence of aerosol GSD, which was varied from 1.5 to 3.5.

Fig. 1. Calculated penetration and deposition efficiencies for log-normally distributed aerosols for mouth breathing with 750 ml tidal volume and 15 breaths/min breathing frequency. Variations in penetration and deposition efficiency about a given MMAD value are due to the influence of aerosol GSD, which was varied from 1.5 to 3.5.
Method induced misclassification generally severe and were greatest for aerosols with MMAD < 1 µm (∼100%) and MMAD 5–15 µm (65–95%). Misclassification rates were moderate (<20%) only for extremely coarse aerosols of MMAD > 15 µm. Misclassification rates for oral and nasal breathing at 750 and 1500 ml tidal volume and 15 breaths/min breathing frequency were similar for each aerosol examined.

**DISCUSSION**

These results are consistent with previous findings by others regarding the substantial exposure misclassification that can occur in epidemiological studies and development of dose–response relationships for exposure standards. Hewett (1991) and Seixas et al. (1995) noted that where there is the potential for workers in a given occupational cohort to be exposed to aerosols having a range of size distributions, such that there is not a constant proportion between penetration and deposition, misclassification may greatly reduce our ability to statistically demonstrate health effects differences between exposure subgroups. Additionally, exposure standards based on respirable aerosol measurement data for one industry, operation...
or even location may not be adequately protective for workers of another industry, operation or location who are exposed to aerosols with substantially different size distributions. For example, the vanadium pentoxide Threshold Limit Value® (ACGIH, 2003) is the same for both dust and fumes, although from these findings one would suspect that workers exposed to fumes at concentrations near the standard would be less adequately protected than if exposed to dust at the same concentration, due to the higher proportion of respirable aerosol that could actually deposit in the pulmonary tissues. In the case of beryllium, exposure to sub-micrometre aerosols might pose a greater risk of beryllium sensitization than would exposure to super-micrometre aerosols at the same concentration (Kent et al., 2001; Paustenbach et al., 2001). Should it be determined that ultrafine particles below 0.1 µm in size represent special hazards, then the 0.1 µm lower bound in our analysis might actually have underestimated the potentially greater risk. As Hewett (1991) demonstrated, even the use of penetration-based respirable aerosol measurements for evaluating relative levels of worker protection may be misleading. He showed that for a substantially sub-micrometre aerosol, changes in the particle size distribution due to engineering controls or process changes providing a large reduction in measured respirable aerosol exposure could actually cause a several-fold increase in deposited dose. In general it might be stated that for a given respirable mass concentration exposure, sub-micrometre aerosols represent a potentially greater threat to workers that super-micrometre aerosols of the same substance.

It seems clear that single cut measurement of respirable dust exposure may not be relied upon for accurate exposure classification in epidemiological studies and standards development. While an ideal solution might appear to be the development of a respirable aerosol sampler that exactly duplicates the Heyder et al. deposition curves, past experience suggests that such a goal is not likely to be achieved any time soon. Indeed, previous efforts to develop samplers that are capable of closely matching the comparatively simple ISO/ACGIH/CEN performance criteria have met with only partial success. One may readily appreciate the greater practicality of performing measurements that characterize the entire aerosol size distribution and then accurately estimate tissue doses through a set of relatively simple calculations. In addition, as previously pointed out by Vincent and Mark (1984), the size distribution data have other uses, such as the potential recalculation of doses at a later data as the understanding of particle deposition and/or particle deposition-related health effects is modified. There are a number of sophisticated approaches, such as personal size distribution-indicating instruments. For example, personal cascade impactor samplers having a sufficiently large number of stages with non-overlapping deposition performance and cut sizes ranging from 0.1 to perhaps 15 µm aerodynamic equivalent diameter might be developed from existing designs such as the Marple personal impactor. If overloading of coarse particle upper stages were to be problematical, parallel stage impactors with varying stage flow rates could prove to be useful if adequate sampling flow rates could be achieved (Esmen and Weyel, 1985; Tuchman et al., 1986). In addition, there are less sophisticated approaches, such as conducting a sufficient number of static particle size distribution measurements. While this approach is conveniently achievable with existing instrumentation, the number of samples necessary to fully characterize the workplace might be inconveniently large.

Little can be done to reinterpret existing respirable air sampling data, unless corresponding particle size distribution information is available. If distribution information is on hand, then tissue dose estimates may be obtained by applying a correction factor to the existing exposure data. This factor may be expressed as the ratio of the mass deposition efficiency to the mass penetration efficiency. Correction factors for aerosols that are either log-normally distributed or may be reduced to two or more constituent log-normally distributed aerosols may be obtained from Table 2. Although the correction factors in Table 2 were developed assuming log-normally distributed aerosols, a correction factor could also be calculated for any aerosol of known size distribution.

Another, perhaps less satisfactory but nevertheless useful, approach might be to estimate the historical size distribution data through a knowledge of

<table>
<thead>
<tr>
<th>MMAD (µm)</th>
<th>Oral breathing</th>
<th>Nasal breathing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>750 ml tidal volume</td>
<td>1500 ml tidal volume</td>
</tr>
<tr>
<td>0.1</td>
<td>0.26</td>
<td>0.29</td>
</tr>
<tr>
<td>0.5</td>
<td>0.14</td>
<td>0.19</td>
</tr>
<tr>
<td>1</td>
<td>0.21</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>0.87</td>
<td>0.57</td>
</tr>
<tr>
<td>10</td>
<td>1.56</td>
<td>0.76</td>
</tr>
</tbody>
</table>
processes. In this case, it is important to point out that such an exercise is wrought with difficulties and pitfalls. Two simple examples of such difficulties may be sited by mentioning the elutriation effect of distance from the source on the particle size distribution and the change in aerosol size distribution with ageing.

CONCLUSIONS

The misclassification problem documented here is very important to epidemiology and must be addressed. It can be argued that when past exposures were high, any reasonable indicator was sufficient to gauge the control measures and reduce the pneumoconiotic disease incidence. However, the reduction in exposures to historically very low levels necessitates a re-visitiation of the exposure measurement methods and potential size distribution-based correction factors.

APPENDIX: A MISCLASSIFICATION METRIC

Let \( m \) subjects be classified in \( N \) ordered classes, \( C_j; i = 1, 2, \ldots, N \). If an arbitrary subject \( S_j; j \in \{1, 2, \ldots, m\} \) is classified as \( S_j \in C_{A_j}; A_j = \{1, 2, \ldots, N\} \) when \( S_j \in C_{T_j}; \) then the metric which measures misclassification is related to the ordinal difference between \( A_j \) and \( T_j; \)

\[
\zeta_j = \frac{|T_j - A_j|}{m(N - 1)}
\]

For \( m \) subjects the sum of the parameter \( \zeta \) is also a metric and may be considered as a measure of misclassification:

\[
\Xi = \sum_{i=1}^{m} \zeta_j = \frac{\sum_{i=1}^{m} |T_j - A_j|}{m(N - 1)}
\]

Clearly, \( 0 \leq \Xi \leq 1 \). \( \Xi = 0 \) implies that the misclassification is 0, i.e. each subject is classified correctly. If \( \Xi = 1 \) implies that the classification is maximally imperfect, i.e. each classification \( A \) is either highest or lowest and each true class \( T \) is exactly the opposite. If we assign \( A \) at random with each class having an equal probability of assignment (uniformly random), then \( E(\Xi) = 1/2 \). Therefore, for any set of classified data, if \( \Xi \geq 1/2 \) then the deliberate classification scheme leads to misclassification levels equal or worse than random assignment of classes and if \( \Xi < 1/2 \) then the deliberate classification is better than random assignment. However, one would like to commit acceptably low levels for misclassification. \( \Xi < 0.20 \) implies that at most 20% of the subjects are misclassified at most by 2 units. It may be argued that any level in excess of this number indicates an unacceptable level of misclassification.

Acknowledgement—This project was supported in part by a grant from the US Environmental Protection Agency (R82-6786-010). The article does not necessarily reflect the views of the Agency and no official endorsement should be inferred.