

The Diesel Exhaust in Miners Study: V. Evaluation of the Exposure Assessment Methods

PATRICIA A. STEWART^{1,3}, ROEL VERMEULEN^{1,4}, JOSEPH B. COBLE^{1,5},
AARON BLAIR¹, PATRICIA SCHLEIFF², JAY H. LUBIN¹,
MIKE ATTFIELD^{2†} and DEBRA T. SILVERMAN^{1†*}

¹*Division of Cancer Epidemiology and Genetics, US National Cancer Institute, Bethesda, MD 20892-7240 USA;* ²*Surveillance Branch, Division of Respiratory Disease Studies, US National Institute for Occupational Safety and Health, Morgantown, WV 26505 USA*

Received 14 June 2010; in final form 31 January 2011; published online 2 March 2012

Exposure to respirable elemental carbon (REC), a component of diesel exhaust (DE), was assessed for an epidemiologic study investigating the association between DE and mortality, particularly from lung cancer, among miners at eight mining facilities from the date of dieselization (1947–1967) through 1997. To provide insight into the quality of the estimates for use in the epidemiologic analyses, several approaches were taken to evaluate the exposure assessment process and the quality of the estimates. An analysis of variance was conducted to evaluate the variability of 1998–2001 REC measurements within and between exposure groups of underground jobs. Estimates for the surface exposure groups were evaluated to determine if the arithmetic means (AMs) of the REC measurements increased with increased proximity to, or use of, diesel-powered equipment, which was the basis on which the surface groups were formed. Estimates of carbon monoxide (CO) (another component of DE) air concentrations in 1976–1977, derived from models developed to predict estimated historical exposures, were compared to 1976–1977 CO measurement data that had not been used in the model development. Alternative sets of estimates were developed to investigate the robustness of various model assumptions. These estimates were based on prediction models using: (i) REC medians rather AMs, (ii) a different CO:REC proportionality than a 1:1 relation, and (iii) 5-year averages of historical CO measurements rather than modeled historical CO measurements and DE-related determinants. The analysis of variance found that in three of the facilities, most of the between-group variability in the underground measurements was explained by the use of job titles. There was relatively little between-group variability in the other facilities. The estimated REC AMs for the surface exposure groups rose overall from 1 to 5 $\mu\text{g m}^{-3}$ as proximity to, and use of, diesel equipment increased. The alternative estimates overall were highly correlated (~ 0.9) with the primary set of estimates. The median of the relative differences between the 1976–1977 CO measurement means and the 1976–1977 estimates for six facilities was 29%. Comparison of estimated CO air concentrations from the facility-specific prediction models with historical CO measurement data found an overall agreement similar to that observed in other epidemiologic studies. Other evaluations of components of the exposure assessment process found moderate to excellent agreement. Thus, the overall evidence suggests that the estimates were likely accurate representations of historical personal exposure levels to DE and are useful for epidemiologic analyses.

Keywords: diesel exhaust; elemental carbon; exposure assessment; mining

*Author to whom correspondence should be addressed. Tel: +301-435-4716; fax: +301-402-1819; e-mail: silvermd@mail.nih.gov

†Co-senior authors.

³Present address: Stewart Exposure Assessments, LLC, Arlington, VA 22207 USA

⁴Present address: Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

⁵Present address: 1412 Harmony Lane, Annapolis, MD 21409 USA

INTRODUCTION

Investigators at the US National Cancer Institute (NCI) and the US National Institute for Occupational Safety and Health (NIOSH) conducted the Diesel Exhaust in Miners Study (DEMS) to examine causes of mortality, particularly from lung cancer, associated with diesel exhaust (DE) among underground miners. A retrospective exposure assessment was conducted to provide estimates of average exposure levels, by mining facility, department, job title, and year to respirable elemental carbon (REC), a component of DE.

This paper describes an evaluation of the exposure assessment process. The goal was to determine the accuracy of the estimates developed for the study and whether the estimates were useful for investigating mortality risk associated with DE exposure.

Several approaches were used in the evaluation because there was no gold standard and because the independent data used in the evaluations varied in type, quantity, and quality by facility. Some information on these evaluations was presented in other reports (Stewart *et al.*, 2010; Vermeulen *et al.*, 2010a,b).

An overview of the exposure assessment steps is presented with an indication where an evaluation was made (Table 1). A description of each evaluation is then presented. Results are presented in order of the evaluation description in the Methods section.

METHODS

Overview of the exposure assessment methods

The primary objective of the exposure assessment process was to develop annual quantitative estimates

Table 1. Summary of the independent data used for the evaluations.

Step in exposure assessment process	Type of evaluation; data evaluated	Comparison data
Processing of work histories	Percent agreement; locations, identified by long-term workers, where study subjects worked who had missing location in work histories	Facility records ($n = 1640$ locations)
Selection of the AM as the exposure metric	Pearson correlation; cumulative exposure levels of underground study subjects from estimates based on AMs of 1998–2001 DEMS REC personal measurements and the primary prediction models	Cumulative exposure levels of underground study subjects from estimates based on medians of 1998–2001 DEMS REC personal measurements and the primary prediction models ($n = 8344$ subjects)
Underground exposure estimates—development of underground exposure groups	Relative difference and Pearson correlation; TWAs based on estimates of proportion of time jobs spent in major underground areas and 1998–2001 DEMS area measurements	1998–2001 DEMS REC personal measurements ($n = 97$ jobs)
Underground exposure estimates—assumptions used for modeling historical trends	Analysis of variance; estimated between- and within-group variability of REC measurements	1998–2001 DEMS REC personal measurements (U1 $n = 97$ jobs; U2 $n = 39$ groups; U3 $n = 22$ groups) ^a
Underground exposure estimates—time trends	Relative difference: REC estimates for underground jobs derived from the primary prediction models using a 1:1 CO:REC proportionality	REC underground estimates derived from Yanowitz <i>et al.</i> (2000) and US EPA (2002) findings of a differential change of DPM to CO over time ($n = 172$ facility-years)
Underground exposure estimates—time trends	Pearson correlation: REC cumulative exposure levels of underground study subjects from estimates derived from the primary prediction models using a 1:1 CO:REC proportionality	Cumulative REC estimates derived from: (i) REC = CO ^{0.58} relation seen in the 1998–2001 DEMS REC and CO area measurements, (ii) 1976–2001 5-year average CO area measurements. Both $n = 8344$ subjects
Underground exposure estimates—time trends	Relative difference: 1976–1977 CO face area estimates derived from the primary prediction models in six facilities	1976–1977 CO short-term face area measurement data $n = 6$ means) (Sutton <i>et al.</i> , 1979)
Underground exposure estimates—time trends	Relative difference: 1994 REC personal estimates in Facility B	1994 REC personal measurements ($n = 2$ jobs) (Stanevich <i>et al.</i> , 1997)
Surface exposure estimates	Monotonic trends: REC personal estimates and % NDs by surface groups	1998–2001 DEMS personal measurements ($n = 24$ exposure groups)

TWA, time-weighted average; DPM, diesel particulate matter; CO, carbon monoxide; ND, non detectable measurement.

^aSee footnote b of Table 3 for definition of U1, U2, and U3.

of DE exposure levels, using REC, for all jobs at eight mining facilities from the first year of diesel use (1947–1967, depending on the facility) to 31 December 1997 for use in the epidemiologic analyses. The procedure was

- REC and carbon monoxide (CO) personal and area measurements were collected during air monitoring surveys in 1998–2001 at seven of the eight study facilities (the DEMS surveys) (Coble *et al.*, 2010; Vermeulen *et al.*, 2010b). (The eighth facility had closed permanently in 1993.) These measurements were, essentially, the only REC measurements available for the study facilities. Work histories were abstracted from the mining facilities' personnel records. Long-term workers provided information on the small number of work locations that were missing from the work history records. The accuracy of this information was assessed (See Jobs with missing location information).
- Exposure groups were developed independently of REC personal or area measurements and separately for underground and for surface jobs based on exposure determinants collected during interviews with long-term workers. These groups then were compared to the REC measurements. (See Underground Exposure Estimates, Development of Underground Exposure Groups; Surface Exposure Estimates).
- The arithmetic means (AMs) of the REC measurements on the subset of jobs measured in the DEMS surveys were assigned as the 1998–2001 reference estimates to all jobs within each exposure group. A sensitivity analysis compared, as the exposure metric, the AM with the median (See REC AMs as the Exposure Metric).
- Historical CO air concentrations at the production face were used as a surrogate for DE exposures and estimated back to the first year of dieselization in the underground locations. The concentrations were modeled using primarily US Mine Safety and Health Administration (MSHA) Mine Information Data System (MIDAS) historical area CO compliance data and DE-related determinants [e.g. engine horsepower adjusted for percentage of time used (ADJ HP) and total air flow rate exhausted from the underground operations in cubic feet per minute (CFM), i.e. ADJ HP/CFM] (1 HP = 0.746 kW; 1 CFM = 1.7 m³ h⁻¹). The results of this modeling are called herein the primary prediction models. Sensitivity analyses were conducted on the assumptions used in the models. (See Under-

ground Exposure Estimates, Assumptions used for Modeling Historical Trends).

- The annual modeled CO air concentrations relative to 1998–2001 CO personal measurements from the DEMS surveys were used to adjust the job-specific 1998–2001 REC reference estimates back to the start of dieselization to provide historical annual REC personal exposure estimates. The final estimates were compared to independent historical measurement data. (See Underground Exposure Estimates, Time Trends).
- REC personal exposure estimates for surface jobs were adjusted for temporal changes only when jobs changed exposure groups (see Stewart *et al.*, 2010) (such as when diesel-powered equipment replaced gasoline-powered equipment). This approach was taken because of the very low exposure levels of workers on the surface compared to those underground (Coble *et al.*, 2010) and because of the less specific information available on surface diesel equipment.
- The estimates for both underground and surface jobs were then combined with individual work histories to derive estimates of personal cumulative exposure and other exposure metrics. The estimates derived from the primary prediction models were considered the primary estimates for the epidemiologic analysis.

Details and other information on the exposure assessment process are available elsewhere (Coble *et al.*, 2010; Stewart *et al.*, 2010; Vermeulen *et al.*, 2010a,b;).

Jobs with missing location information

About 5% of the job entries from the personnel records lacked information on whether the location of the job was on the surface or underground. It was critical to obtain this information because of the large differences in exposure levels between surface and underground workers (Coble *et al.*, 2010). Work histories of individuals with one or more jobs with an unknown location lasting at least 2 years were reviewed with long-term employees at the individuals' mining facility to determine the work location. After the site visits had been completed, the work histories for two facilities that had particularly complex work history records were re-abstracted to ensure high quality data. In addition, for these and the other mining facilities, seniority records and other information in the personnel records not originally abstracted were reviewed for location and used to replace the missing information. For this evaluation, we calculated the percent of overall agreement for the

locations identified during the interviews and the locations from the additional information in the personnel records, overall and by facility, by surface and underground classification.

REC AMs as the exposure metric

The AM was used in the exposure assessment process because AMs are considered the best statistic for calculating cumulative exposure in evaluations of chronic disease (Seixas *et al.*, 1991). Accordingly, the 1998–2001 job-specific REC reference estimates were based on the AM of the DEMS REC full-shift personal measurements (Coble *et al.*, 2010). To investigate the robustness of the AM to high or low measurements, a set of alternate REC exposure estimates was developed for all underground jobs by replacing the 1998–2001 REC AMs by the 1998–2001 medians of the REC measurements and using the relative time trends derived from the primary prediction models. Cumulative exposure levels from both sets of estimates were calculated for all underground study subjects. Pearson correlation coefficients (r_P) were calculated overall and by mining facility.

Underground exposure estimates

Development of underground exposure groups. Because of limited job-specific REC measurement data, we initially developed three types of exposure groups for each facility's underground operation using the following criteria in order of decreasing specificity (homogeneity): U1, in which each group, represented by a single standardized job title, comprised all job entry titles from the personnel records that differed only in spelling or word order and that facility personnel had identified as having performed the same job tasks; U2, sets of groups, each comprising several U1 groups whose workers spent similar proportions of time (as indicated in the interviews) in four major areas of the underground operations (i.e. the production face, haulage and travel ways, shop and office area, and crusher area); and U3, sets of U2 groups whose members spent time in areas with similar CO air concentrations based on the historical MIDAS CO area measurements (Coble *et al.*, 2010). All groups were developed without regard to REC measurements. [Two other types of groups (contributing 14% of the person-years) were developed but could not be evaluated; see Stewart *et al.*, 2010]. Once the groups were developed, the REC full-shift personal measurements from the DEMS survey were assigned to each group. Each job title was then assigned and blinded to the REC measurement levels, to one of the three exposure groups,

based on its most specific group having at least five REC measurements. The AM of the REC personal measurements assigned to each group became the reference 1998–2001 exposure estimate.

As indicated above, the grouping of jobs to the U2 and, indirectly, the U3 exposure groups was dependent on the proportion of time workers in the various jobs spent in four major underground areas that differed substantially with regard to DE concentrations (Coble *et al.*, 2010). This assignment of time and area to each job could be tested by using those same proportions of times and the REC full-shift area measurements collected in those areas during the DEMS surveys (Coble *et al.*, 2010). First, the AMs of the REC area measurements taken in each of the four areas were calculated for each facility (Coble *et al.*, 2010). For each job measured in the DEMS survey, each area AM was weighted by the respective proportion of a work shift that workers in the job spent in each area and summed across all areas to develop an 8-h time-weighted average (TWA). Differences and relative differences (the difference divided by the observed mean) were calculated between the observed AMs, i.e. of the full-shift personal measurements, and the predicted TWAs. Pearson correlation coefficients were also calculated between the predicted and observed values of each job by facility and overall.

To determine how much variability in the REC personal measurements was explained by the three underground exposure groups (i.e. U1–U3), variance component models were constructed for each facility to estimate the between- and within-group variability of the measurements. For U1 groups (i.e. the standardized job titles), a random effects model was used. For U2 groups (time spent in the four underground areas) and U3 groups (time spent in areas with similar CO air concentrations), mixed effects models were constructed, where the U2 or U3 groups were introduced into the model as fixed effects and the jobs within the exposure group were treated as random effects.

Assumptions used for modeling historical trends. To develop historical estimates for underground jobs, reference REC personal exposure estimates were adjusted using facility-specific CO time trend prediction models. The models used the assumptions that a change in CO levels resulted in an equal change in REC levels and that this relationship was constant over time up to the early 1990s, i.e. a 1:1 proportionality (Vermeulen *et al.*, 2010a). (Adjustment to this proportionality to account for cleaner engines after 1990 was made in the model but is not applicable to this discussion; see Vermeulen

et al., 2010a.) Diesel engines underwent technologic changes before the 1990s, however, and these changes affected emissions of CO and diesel particulate matter (DPM, of which the primary component is REC (Ramachandran and Watts, 2003; Noll *et al.*, 2007). The US Environmental Protection Agency (US EPA, 2002) modeled data generated by Yanowitz *et al.* (2000) on emissions of diesel engines manufactured between 1976 and the late 1990s. The US EPA found that CO and DPM emissions increased with the year of manufacture at only a slightly different rate (annual parameter estimate difference = -0.003 , i.e. that DPM emissions increased slightly less than CO emissions for each earlier manufacturing year). To estimate the effect of this finding on the primary model estimates, we calculated the mean purchase year of all diesel equipment in use underground using for each piece, the purchase year weighted by its ADJ HP. This procedure was done for each year for each facility from 1997 back to 1976. The effect of the two sets of assumptions for each facility was calculated by comparing the 1976 REC estimates derived from the primary models (i.e. with the 1:1 proportionality) and the 1976 REC estimates adjusted for the parameter estimate of -0.003 . The difference is the approximate likely difference in our estimates as a result of changes in engine technology not already accounted for in the primary estimates.

In addition, a sensitivity analysis was conducted based on the relation of $REC = CO^{0.58}$ seen in the DEMS area measurements (Vermeulen *et al.*, 2010a,b). An alternate set of historical prediction trends for underground jobs was developed for each facility using this relation, rather than the 1:1 proportionality used in the primary models. This set of models was called the $REC = CO^{0.58}$ models. A second sensitivity analysis evaluated a set of time trend estimates developed from the observed 5-year averages of the MSHA MIDAS historical CO measurements used for the prediction models, without the DE-related determinants. For years prior to 1976, when there were no CO measurements, the ratio of ADJ HP/CFM to the 1976 ADJ HP/CFM was used to adjust the CO averages (Vermeulen *et al.*, 2010a). These models were called the 5-year average CO models. A set of cumulative estimates for underground workers was calculated using the $REC = CO^{0.58}$ models, as was a second set using the 5-year average CO models. Pearson correlations were calculated between the cumulative REC exposure levels derived from the primary estimates and the cumulative exposure levels derived from the two alternate sets of estimates.

Time trends. Measurement data from an US Mine Enforcement and Safety Administration (MESA)/US Bureau of Mines (BoM) monitoring study conducted in 1976–1977 in six of the study facilities (Sutton *et al.*, 1979) were not used in the development of the primary prediction models and so could be used for method evaluation (Vermeulen *et al.*, 2010a). CO area measurements at the face were selected for comparison because the primary models estimated CO face concentrations and CO personal measurements were not available. Thus, the facility-specific AMs of the 1976–1977 face measurements were compared to the 1976–1977 CO estimates derived from the facility-specific primary models. A second comparison was made of the AMs of the REC full-shift personal measurements taken in 1994 at one mining facility (B) (Stanevich *et al.*, 1997) to the 1994 REC estimates derived from the primary model for two of four underground jobs for which the data could be matched. For both the CO and REC comparisons, differences and relative differences were calculated between the predicted and observed values.

Surface exposure estimates

Interviews with long-term workers provided information on the frequency of use of, and proximity to, diesel-powered equipment. The information was used to categorize jobs into one of three exposure groups: (i) jobs in which workers had no or very limited contact with diesel equipment (Exposure group A), (ii) jobs in which workers drove a diesel forklift truck indoors or operated heavy equipment (>75 HP) <4 h per shift on average, drove light diesel equipment (≤ 75 HP), or worked in close proximity to diesel-powered equipment on a regular basis (Exposure group B), and (iii) jobs in which workers operated heavy diesel equipment or drove a diesel forklift truck indoors for ≥ 4 h per shift on average, and mechanics and maintenance workers who repaired diesel equipment (Exposure group C). The DEMS REC personal measurements were assigned to each of the A, B, and C groups within each facility (S1 estimation groups); across all facilities of the same ore type (i.e. potash, trona, salt, or limestone) (S2 estimation groups); and across all facilities (S3 estimation groups) because of the limited numbers of measurements. The assignments were made blinded to the measurement levels. Each job was assigned the AM of its surface exposure group (A–C) at the most specific facility level (S1–S3) having at least five REC personal measurements.

To evaluate the surface groupings, the AM of each exposure group's REC measurements and the

percentage of nondetectable (ND) measurements were calculated to determine if increasing REC levels, and decreasing percentages of ND measurements, were associated with increasing proximity to, or use of, diesel equipment.

Only Pearson correlation coefficients are presented in this report because the Spearman correlation coefficients were similar. All elements of the exposure assessment process, including this evaluation work, were done blind to the disease outcomes of the epidemiologic studies.

RESULTS

Jobs with missing location information

For the 1640 jobs missing location data, there was 93% agreement (ranging from 86 to 100% by facility) between the location ascertained from the facility records and from the long-term workers. The agreement was similar for jobs identified in the records as being on the surface or as being underground (Supplementary data are available at *Annals of Occupational Hygiene* online).

REC AMs as the exposure metric

A sensitivity analysis compared the use of REC medians in the estimation of cumulative exposure levels of underground workers to cumulative levels based on the REC AMs that were used for the computation of cumulative exposure levels in the epidemiologic analysis. There was an excellent corre-

lation between the two sets of cumulative exposure levels overall ($r_P = 0.98$) and by facility ($r_P = 0.98$ to >0.99) (Supplementary data are available at *Annals of Occupational Hygiene* online).

Underground exposure estimates

Development of underground exposure groups. A TWA was derived from the 1998 to 2001 REC area measurements and estimates of the proportion of time spent in the four underground areas for each job with 1998–2001 personal measurements and compared to the corresponding REC full-shift personal AM that was used as the 1998–2001 reference value. The TWAs overestimated the full-shift measurements by -19% (-48 to 20% by facility) (Table 2). The median correlation was 0.83 (0.15–0.72 by facility).

The analysis of variance results on the underground job groups' REC measurements indicated that the U1 set of groups (i.e. standardized job titles) explained more between-group variability than the U2 and U3 sets of groups in three facilities (Facilities B, E, and G), but they explained little of the variance for the other facilities (Table 3). There was little difference in the within-job or -group variance across the three sets of groups for any facility. For example, Facility B's between-group variability was 0.12 for its U1 groups and 0.00 for its U2 and U3 groups, whereas the within-group variability was essentially the same for the three sets of groups (i.e. 3.28, 3.24, and 3.27, respectively). No change to the exposure group assignments was made as a result of this evaluation.

Table 2. Comparison of estimated REC 8-h TWA^a personal levels to DEMS REC full-shift personal measurements for underground jobs.

Facility ^b	<i>N</i> Jobs	AM of the DEMS REC personal measurement AMs ($\mu\text{g m}^{-3}$)	AM of estimated REC TWAs ($\mu\text{g m}^{-3}$)	Difference ($\mu\text{g m}^{-3}$)	Relative difference ^c (%)	r_P
A	14	339	472	-133	-39	0.39
B	9	107	158	-52	-48	0.52
D	14	109	88	22	20	0.15
E	17	65	88	-23	-35	0.72
G	11	38	38	0	1	0.68
H	11	84	108	-24	-29	0.62
I	21	63	69	-7	-11	0.21
All	97	113	135	-22	-19	0.83

N, number of jobs; r_P , Pearson correlation coefficient.

^aEstimated REC TWA (in micrograms per cubic metre) was calculated, by facility and overall as: [area concentration_{face} × (% time_{face}/100)] + [area concentration_{haulage/travel ways} × (% time_{haulage/travel ways}/100)] + (area concentration_{shop & office} × (% time_{shop & office}/100)) + [area concentration_{crusher} × (% time_{crusher}/100)].

^bMeasurements in Facility J were not available because the facility had closed prior to the monitoring.

^cRelative difference was calculated by dividing the difference by the corresponding observed AM. The differences and relative differences may not be exact because of rounding.

Table 3. Variance components for log-transformed DEMS REC personal measurements of underground exposure groups, by facility.

Facility ^a	U1 exposure groups ^b			U2 exposure groups			U3 exposure groups		
	<i>N</i>	S_{BJ}^2	S_{WJ}^2 ^c	<i>N</i>	S_{BG}^2	S_{WG}^2	<i>N</i>	S_{BG}^2	S_{WG}^2
A	13	<0.01	0.24	5	0.03	0.24	3	<0.01	0.24
B	9	0.12	3.28	5	0.00	3.24	4	0.00	3.27
D	8	0.16	0.54	5	0.35	0.46	3	0.54	0.46
E	16	0.48	0.81	7	0.15	0.80	3	0.16	0.73
G	13	0.35	0.72	5	0.10	0.70	3	0.15	0.69
H	16	0.00	1.21	6	0.00	1.23	3	0.00	1.22
I	21	0.08	0.73	6	0.07	0.75	3	0.07	0.73

N, number of groups in each facility; S_{BJ}^2 , between-job variance; S_{WJ}^2 , within-job variance; S_{BG}^2 , between-group variance; S_{WG}^2 , within-group variance.

^aMeasurements in facility J were not taken because the facility had closed prior to the monitoring.

^bU1 exposure groups were standardized job titles of underground jobs and covered 40% of the person-years of the study. U2 exposure groups comprised sets of standardized job titles based on the percent of time workers with these jobs worked in each of four underground areas (the production face, haulage or travel ways, maintenance shop and offices and, where existent, crusher area) (40% of the person-years). U3 exposure groups comprised sets of U2 groups, based on similar CO area air concentrations (6% of the person-years). The two remaining grouping methods, representing 14% of the person-years, could not be evaluated.

^cIncludes within-day variability.

Assumptions used for modeling historical trends. There was little difference between the primary REC estimates for underground jobs in 1976 and the REC estimates based on the changing relationship of DPM and CO in engine emissions from the 1990s back to 1976, as published by Yanowitz *et al.* (2000) and the US EPA (2002). The 1976 REC estimates would have been only 10% lower than estimated in the study had the Yanowitz *et al.* (2000) data represented the trends in engine emissions in our facilities.

The sensitivity analysis exploring the robustness of the assumptions used in the time trend prediction models generally found very high correlations between the cumulative exposure levels of underground workers using alternate assumptions with those derived from the primary models. The results from the model using the $REC = CO^{0.58}$ relationship led to an overall correlation of 0.88 ($r_p = 0.96$ – 0.99 by facility) (Supplementary data are available at *Annals of Occupational Hygiene* online). The 5-year average CO estimates (i.e. using only the measurement data after 1975 and the ADJ/CFM ratio before that) resulted in an overall correlation with the primary estimates of 0.87 ($r_p = 0.95$ to >0.99 by facility).

Time trends. Comparison between observed underground 1976–1977 CO face measurement AMs and the underground 1976–1977 CO face concentrations estimated from the primary prediction models found that the median of the relative differences between the two sets of estimates was 29% (–25 to 49% by facility) (Table 4). Figure 1 displays the

1976–1977 CO estimates and the variability of the 1976–1977 CO measurements. Comparison of the 1994 REC estimates developed from the primary models to 1994 REC measurements in Facility B found that the primary estimate was higher than the measurement mean for the continuous miner by 10% and for the foreman by 6%.

Surface exposure estimates

Overall, the estimated REC levels for Surface exposure groups A, B, and C increased with increasing contact with DE (1, 3, and 5 $\mu\text{g m}^{-3}$, respectively) (Table 5). The AMs also increased monotonically across the exposure groups in five facilities (Facilities B, D, H, I, and J). In Facility A, the estimate for exposure group A was overridden due to a relatively higher concentration when compared with the estimate for Exposure group B in that facility or for Exposure groups A or B in any other facility. For all facilities combined, the percentage of NDs decreased as surface exposure groups potentially had more contact with DE (75, 57, and 47% respectively, for Exposure groups A, B, and C).

DISCUSSION

Evaluating exposure assessment methods is an important component of epidemiologic studies. Studies, however, often lack the data to make such evaluations. In this study, we collected information that could be used to evaluate the exposure assessment process because we had no gold standard. We

Table 4. Comparison of underground estimated CO production face area and REC personal levels to independent CO production face area and REC personal measurement AMs, by facility and overall.

	Measurements		Estimates	Difference	Relative difference ^a (%)
	N	AM			
1976–1977 CO area (p.p.m.) ^b					
Facility ^c					
B	90	7.2	5.2	2.1	29
D	136	10.5	8.0	2.5	24
E	148	8.5	10.6	−2.1	−25
H	100	7.7	3.9	3.8	49
I	122	7.7	4.9	2.9	37
J	217	8.1	4.4	3.7	46
All	813	8.3	6.1	2.2	29 ^d
1994 REC personal ($\mu\text{g m}^{-3}$) ^e					
Job title					
Continuous miner	26	248	273	25	−10
Foreman	6	166	176	10	−6

N, number of measurements; p.p.m., parts per million.

^aRelative difference was calculated by dividing the difference by the corresponding observed AM. Differences or relative differences may not be exact because of rounding.

^bFrom Sutton *et al.* (1979).

^c1976–1977 measurement data were not available for facilities A and G.

^dMedian of the differences.

^eFrom Stanevich *et al.* (1997).

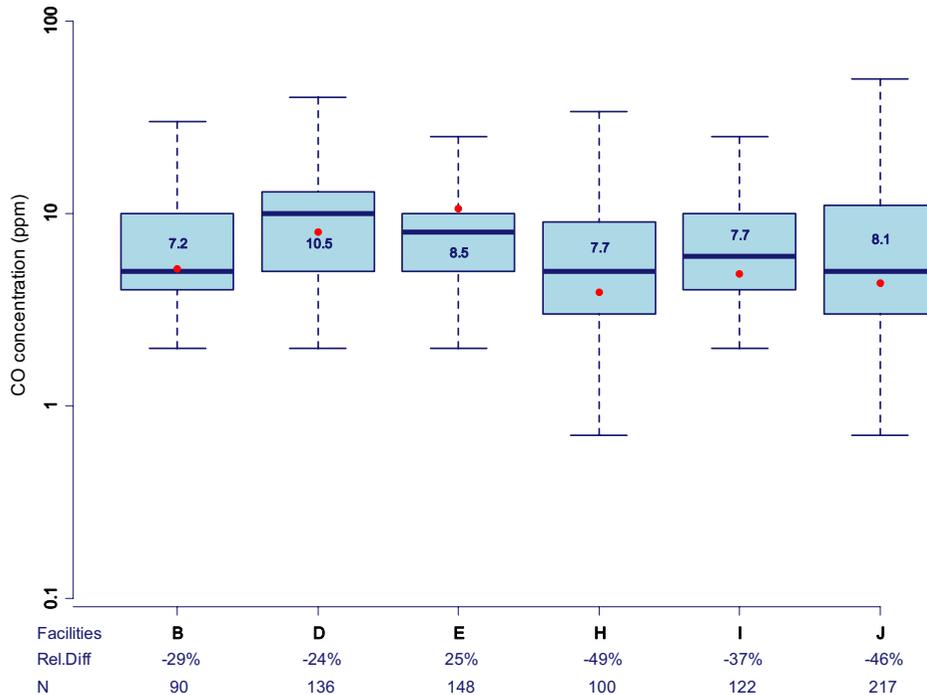


Fig. 1. CO, Carbon monoxide; Rel. Diff, Relative difference. For the definition, see footnote a in Table 4; N, Number of CO measurements from 1976–1977 MESA/BoM survey Sutton *et al.*, (1979). The boxes display the 25th and 75th percentiles, the horizontal line within each box displays the median, and the number plotted represents the mean of the MESA/BoM survey. The vertical whiskers extend 1.5 times the interquartile range above and below the boxes. The dots indicate the predicted 1976–1977 CO estimates by mining facility.

Table 5. Estimated REC personal exposure levels and percentage of nondetectable levels from the DEMS REC personal measurement data, by facility and surface exposure category.

Facility	Surface exposure group A ^a		Surface exposure group B		Surface exposure group C	
	Estimated REC level ^b ($\mu\text{g m}^{-3}$) (<i>N</i>)	% ND	Estimated REC level ($\mu\text{g m}^{-3}$) (<i>N</i>)	% ND	Estimated REC level ($\mu\text{g m}^{-3}$) (<i>N</i>)	% ND
A	2 ^c (96)	75	4 (11)	9	11 (6)	0
B	1 (9)	56	3 (50)	34	5 ^d (34)	47
D	1 (20)	95	2 (12)	83	5 ^d (34)	47
E	1 (6)	100	4 (11)	64	4 (8)	75
G	2 (12)	75	2 (17)	94	4 ^c (17)	53
H	1 (8)	88	2 (22)	77	4 ^c (17)	53
I	1 (25)	92	2 (10)	80	5 (12)	42
J ^f	1 (29)	83	3 (62)	44	5 ^d (34)	47
All ^g	1	75	3	57	5	47

N, number of measurements; %ND, percent of nondetectable measurements.

^aSurface exposure category definitions—A: jobs in which workers had no or very limited contact with diesel equipment; B: jobs in which workers drove a diesel forklift truck indoors or operated heavy equipment (>75 HP) <4 h per shift on average, drove light diesel equipment (≤ 75 HP), or worked in close proximity to diesel-powered equipment on a regular basis; and C: jobs in which workers operated heavy diesel equipment or drove a diesel forklift truck indoors for ≥ 4 h per shift on average, and mechanics and maintenance workers who repaired diesel equipment.

^bFor the calculation of the AM, the concentrations for the ND measurements were imputed based on a distributional Maximum Likelihood Estimation procedure (Helsel, 2005; Vermeulen *et al.*, 2010a).

^cThe AM of the measurements ($\text{AM} = 5 \mu\text{g m}^{-3}$; $n = 6$) for this exposure category was higher than the AM for Surface category B in this facility ($\text{AM} = 4 \mu\text{g m}^{-3}$; $n = 11$) and two to five times higher than the same surface category in any other facility. Measurements for this category were derived from the AM of Surface category A measurements, pooled across all facilities.

^dThe AM was derived from the AM of Surface category C measurements pooled across all facilities because there were fewer than five measurements for this category in each facility and in all potash facilities combined.

^eThe AM was derived from the AM of Surface category C measurements pooled across facilities of the same ore type (i.e. trona) (G, H, and I) because there were fewer than five measurements for this category in Facilities G and H.

^fThe AMs for Facility J were derived by pooling the measurements for Facilities B and D for Exposure category A and B.

^gThe number of measurements on which the estimated AMs were based does not add up to the total number of measurements because of the grouping strategy when there were fewer than five measurements per facility/surface category. See text for details.

compared DE-related historical exposure estimates to DE-related measurements, i.e. 1976–1977 CO measurement data and 1994 REC measurement data. We compared our estimates to recent data (1998–2001 REC area measurements to 1998–2001 REC personal measurements and trends in 1998–2001 REC personal measurements by surface category). We also conducted sensitivity analyses to evaluate our assumptions. Despite the inherent limitations in the data and the comparisons, we found moderate to excellent agreement between our estimates and these other data sources.

A critical evaluation is the comparison of our 1976–1977 CO estimates predicted for the face to the CO measurement data collected in 1976–1977 at the face (Sutton *et al.*, 1979). We found little overall difference between the two sets of data: the median of the differences between the estimates and the observed measurements was 29% and ranged from –25 to 49% by facility (Vermeulen *et al.*, 2010a). In addition to the CO comparisons, two underground jobs with 1994 REC estimates could be

directly matched to the measured REC in 1994 in Facility B (Stanevich *et al.*, 1997). The REC estimates were higher than the REC measurements by $\leq 10\%$.

When we compared these differences to measurement error reported in the literature, we found similar levels of agreement. In a study of side-by-side acrylonitrile measurements, in which two sets of measurements were taken concurrently on the same individuals using the same sampling and analytic method, mean differences of –14 to 21% were found, depending on the sampling method (Zey *et al.*, 2002). A study of elemental carbon measurements taken side-by-side with our measurements generally found mean differences of –18 to 40% by facility and job function (Cohen *et al.*, 2002) (Supplementary data are available at *Annals of Occupational Hygiene* online). Thus, the relative differences between the estimates and the measurement data were in the range of variability that has been seen in actual measurement studies. These results indicate that the estimates were reasonably accurate and support the use of the MSHA area CO

compliance measurement data, estimates of ADJ HP and other determinants to predict relative trends in DE exposure levels.

Grouping of jobs is common in epidemiologic studies because measurement data are often sparse. As has been done by others (Quinn *et al.*, 2001), we used job elements (which also could be called exposure determinants, e.g. percent time of a work shift spent in various underground areas) based on information obtained by interview to group jobs with and without DEMS REC measurements. We asked the workers to identify historical changes to ensure that our data reflected the standard procedures followed over time. Only a few jobs were reported as having changed and, in these cases, a new job title was developed to allow a different set of estimates. When we compared estimated TWAs, calculated from the DEMS area measurements and proportion of time in these areas, with full-shift DEMS personal measurements, the relative difference was -19% . This difference is similar to that found for side-by-side measurements (Cohen *et al.*, 2002; Zey *et al.*, 2002). The overall correlation between the two sets of data was 0.83. The two facilities with the low correlations (A and I) were likely due to the small number of measurements (<5) for several of the jobs and areas used in the analysis and the homogeneity of exposure levels in these facilities. These findings are important because the basis for assignment of a substantial number of jobs to exposure groups was the proportions of time spent in the various underground areas. The results suggest that the time estimates were accurate enough to rank the jobs similarly to the measurement data.

The goal of any grouping strategy is to maximize between-job (or group) variability and minimize within-job (or group) variability. Grouping of jobs is, therefore, a trade-off between reliability and specificity. We chose to use facility-specific standardized job titles as the preferred method over cruder ways of grouping of underground jobs (e.g. departments, facility, type of product) to increase specificity but used a criterion of at least five measurements per exposure group in order to increase the reliability of the means. The analysis of variance on the DEMS survey data indicated that the selection of this approach over the two broader groups was appropriate in three facilities; in three others, it did not make a difference; and in only one facility was either of the other two types of groups preferable. This facility, however, had the lowest percentage of exposure-years linked to U1 groups (Stewart *et al.*, 2010). The lack of contrast among jobs in some of the mines was not so much a limitation of our job grouping strategy as it

was a reflection of the homogeneity of exposure levels across jobs within the underground operations. Thus, these data support our use of standardized job titles as the preferred grouping strategy. Using the broader groups to represent jobs of possibly different exposure levels was likely to have increased random error in the assigned exposure levels and to have decreased between-job variability, but some type of broader group was necessary for the jobs with no measurements. Thus, some attenuation of the lung cancer risk estimates may have occurred because of the grouping procedure.

The findings from the analyses of variance, i.e. that there were differences in underground jobs by facility as well as the findings of the differences among the facilities seen both with the personal measurements (Coble *et al.*, 2010) and with the area measurements (Coble *et al.*, 2010; Vermeulen *et al.*, 2010b), support our development of facility-specific estimates. We applied the same procedure to all study facilities for consistency and because of the substantial differences in horsepower, ventilation levels, and other environmental factors between the underground operations (Coble *et al.*, 2010; Stewart *et al.*, 2010).

The overall correlation of the cumulative exposure levels of underground workers using the medians with those using AMs was 0.98, with similar findings for the individual facilities. In part, these high correlations are due to the use of the same duration (i.e. number of years worked at a job) values used in both calculations of cumulative exposure. Nevertheless, these correlations suggest that using the AMs, with the occasional inclusion of extreme values, yielded almost the same ranking of study subjects as using medians, which are less affected by extreme values.

Two sets of cumulative exposure estimates were developed from alternate prediction models to evaluate the robustness of the study assumptions used to develop the primary historical REC estimates for underground jobs. The correlation coefficient between the subjects' cumulative REC estimates using the primary historical estimates and those using the power relationship of $REC = CO^{0.58}$ was 0.88. The coefficient comparing cumulative estimates based on the primary estimates with those based only on CO measurement data after 1975 was 0.87. Although the similarity of the correlations was influenced by the use of the same duration values in the calculation of cumulative exposure, these results suggest that the rankings of the study subjects by cumulative exposure were probably not significantly affected by the deterministic assumptions (e.g. ADJ HP/CFM) or the proportionality assumption

(change in CO:change in REC = 1:1) on which the primary models were based.

The surface REC estimates were based on the likely contact with diesel equipment within the various jobs, and assignment of jobs to surface groups was made blinded to the measurement levels. We found an increasing trend in the estimated concentration relative to increasing proximity to diesel equipment for the measured jobs across all facilities and for several individual facilities. Discrepancies away from the expected monotonic trends of REC levels based on diesel contact were generally small (<20%) and, in part, may have been due to small numbers of measurements. The impact on the epidemiologic study of these discrepancies is likely to be minimal because the absolute differences among the discrepant surface workers' exposure estimates was small (<5 $\mu\text{g m}^{-3}$) compared to the mean estimates of the underground workers (78–216 $\mu\text{g m}^{-3}$ by facility).

SUMMARY

Several approaches were used to assess the DE exposure estimates developed for our epidemiologic study of underground miners. Direct comparison of historical CO face measurement data with estimated CO face concentrations from our facility-specific prediction models found excellent overall agreement that was generally within the limits of the sampling variation for measurements seen by others. Other evaluations of components of the exposure assessment process (identification of unknown location in the work histories, the use of arithmetic means to describe exposure levels, development of exposure groups of underground jobs based on estimates of the amount of time workers spent at specific underground areas, the use of a different REC:CO proportionality due to changing engine technologies, the development of alternate estimates based on analyses of recent REC and CO measurements and based on historical CO measurements, and development of surface exposure groups based on contact with diesel equipment) indicated moderate to excellent levels of agreement. Thus, the overall evidence from these evaluations suggests that our REC estimates are likely to be accurate representations of historical personal exposure levels to diesel exhaust that are useful for an epidemiologic study.

SUPPLEMENTARY DATA

Supplementary data can be found at <http://annhyg.oxfordjournals.org/>.

FUNDING

Division of Cancer Epidemiology and Genetics of the National Cancer Institute.

Acknowledgements—We thank the management and employees of the facilities and representatives of the labor unions who participated in this study. Without their help and the extra efforts they made to provide us with historical reports this evaluation would not have been possible. We also thank Dr Noah Seixas at University of Washington for his valuable comments, Rebecca Stanevich and Daniel Yereb (formerly of NIOSH) and Dr Mustafa Dosemeci (formerly of NCI) for their work on the DEMS surveys, Nathan Appel of IMS, Inc. for programming support, and Lutzen Portengen, Utrecht University, for statistical support. The findings and conclusions in this report/presentation have not been formally disseminated by the NIOSH and should not be construed to represent any agency determination or policy.

REFERENCES

- Coble JB, Stewart PA, Vermeulen R *et al.* (2010) The diesel exhaust in miners study: II. Exposure monitoring surveys and development of exposure groups. *Ann Occup Hyg*; 54: 747–61.
- Cohen HJ, Borak J, Hall T *et al.* (2002) Exposure of miners to diesel exhaust particulates in underground nonmetal mines. *Am Ind Hyg Assoc J*; 63: 651–58.
- Helsel DR. (2005) More than obvious: better methods for interpreting nondetect data. *Environ Sci Technol*; 39: 419A–23A.
- Noll JD, Bugarski AD, Patts LD *et al.* (2007) Relationship between elemental carbon, total carbon, and diesel particulate matter in several underground metal/non-metal mines. *Environ Sci Technol*; 41: 710–16.
- Quinn MM, Smith TJ, Youk AO *et al.* (2001) Historical cohort study of US man-made vitreous fiber production workers: VIII. Exposure-specific job analysis. *J Occup Environ Med*; 43: 824–34.
- Ramachandran G, Watts WF Jr. (2003) Statistical comparison of diesel particulate matter measurement methods. *Am Ind Hyg Assoc J*; 64: 329–37.
- Seixas NS, Moulton LH, Robins TG *et al.* (1991) Estimation of cumulative exposures for the National Study of coal workers' pneumoconiosis. *Appl Occup Environ Hyg*; 6: 1034–41.
- Stanevich RS, Hintz P, Yereb D *et al.* (1997) Elemental carbon levels at a potash mine. *Appl Occup Environ Hyg*; 12: 1009–12.
- Stewart P, Coble J, Vermeulen R *et al.* (2010) The Diesel Exhaust in Miners Study: I. overview of the exposure assessment process. *Ann Occup Hyg*; 54: 728–46.
- Sutton GW, Weems GW, Schutz LA *et al.* (1979) Summary report of the environmental results of the MSHA and NIOSH silica/diesel exhaust study. Industrial hygiene for mining and tunneling: proceedings of a topical symposium, November 6 and 7, 1978. Denver, Colorado. Cincinnati, OH: American Conference of Governmental Industrial Hygienists, Inc.
- US EPA. (2002) Health assessment document for diesel engine exhaust. National Center for Environmental Assessment EPA/600/8–90/057F. Washington, DC: US EPA.
- Vermeulen R, Coble JB, Lubin J *et al.* (2010a) The diesel exhaust in miners study: IV. Estimating historical exposures to

- diesel exhaust in underground non-metal mining facilities. *Ann Occup Hyg*; 54: 774–88.
- Vermeulen R, Coble JB, Yereb D *et al.* (2010b) The Diesel Exhaust in Miners Study: III. interrelations between respirable elemental carbon and gaseous and particulate components of diesel exhaust derived from area sampling in underground non-metal mining facilities. *Ann Occup Hyg*; 54: 762–73.
- Yanowitz J, McCormick RL, Graboski MS. (2000) In-use emissions from heavy-duty vehicle emissions. *Environ Sci Technol*; 34: 729–40.
- Zey JN, Stewart PA, Hornung R *et al.* (2002) Evaluation of side-by-side pairs of acrylonitrile personal air samples collected using different sampling techniques. *Appl Occup Environ Hyg*; 17: 88–95.