A critical review of the relationship between occupational exposure to diesel emissions and lung cancer risk

Matthias Möhner & Andrea Wendt

To cite this article: Matthias Möhner & Andrea Wendt (2017): A critical review of the relationship between occupational exposure to diesel emissions and lung cancer risk, Critical Reviews in Toxicology, DOI: 10.1080/10408444.2016.1266598

To link to this article: http://dx.doi.org/10.1080/10408444.2016.1266598

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 09 Feb 2017.

Submit your article to this journal

View related articles

View Crossmark data
ABSTRACT
In 2012, a working group of the International Agency for Research on Cancer classified diesel exhaust (DE) as a human carcinogen (Group 1). This decision was primarily based on the findings of the Diesel Exhaust in Miners Study (DEMS). The disparity between the results of various methodological approaches applied to the DEMS led to several critical commentaries. An expert panel was subsequently set up by the Health Effects Institute to evaluate the DEMS results, together with a large study in the trucking industry. The panel concluded that both studies provided a useful basis for quantitative risk assessments (QRAs) of DE exposure. However, the results of both studies were non-definitive as the studies suffer from several methodological shortcomings. We conducted a critical review of the studies used by the International Agency for Research on Cancer (IARC) working group to evaluate the relationship between DE and lung cancer. The aim was to assess whether the available studies support the statement of a causal relationship and, secondarily if they could be used for QRA. Our review highlights several methodological flaws in the studies, amongst them overadjustment bias, selection bias, and confounding bias. The conclusion from our review is that the currently published studies provide little evidence for a definite causal link between DE exposure and lung cancer risk. Based on two studies in miners, the DEMS and the German Potash Miners study, QRA may be conducted. However, the DEMS data should be reanalyzed in advance to avoid bias that affects the presently published risk estimates.

Abbreviations: AIC: Akaike’s Information Criterion; BIC: Bayesian Information Criterion; BMI: Body Mass Index; CI: Confidence Interval; CO: Carbon Monoxide; DE: Diesel Exhaust; DEMS: Diesel Exhaust in Miners Study; EC: Elemental Carbon; ETS: Environmental Tobacco Smoke; HHE: Healthy-Hire Effect; HR: Hazard Ratio; HWE: Healthy-Worker Effect; HWSE: Healthy-Worker Survivor Effect; IRR: Incidence Rate Ratio; JEM: Job Exposure Matrix; MRR: Mortality Rate Ratio; OR: Odds Ratio; QRA: Quantitative Risk Assessment; RDD: Random Digit Dialing; REC: Respirable Elemental Carbon; RR: Relative Risk; SD: Standard Deviation; SER: Standardized Employment Ratio; SES: Socio-Economic Status; SMR: Standardized Mortality Ratio; TC: Total Carbon; TSFE: Time Since First Exposure; TSH: Time Since Hire

Table of contents
Introduction ................................................................. 2
Methods ................................................................. 3
Study selection .......................................................... 3
Criteria for the methodological evaluation ......................... 3
Statistical methods ..................................................... 3
Results ................................................................. 3
Diesel Exhaust in Miners Study ....................................... 5
Description of the study and authors’ main results ... 5
Comments on the study ............................................. 6
German Potash Miners Cohort Study ................................ 9
Description of the study and authors’ main results ... 9
Comments on the study ............................................. 9
Cohort study in Canadian retired railway workers ... 10
Description of the study and authors’ main results ... 10
Comments on the study ............................................. 10
Cohort study among Finnish locomotive drivers ................. 10
Description of the study and authors’ main results ... 10
Comments on the study ............................................. 10
Cohort study among US railroad workers ....................... 11
Description of the study and authors’ main results ... 11
Comments on the study ............................................. 11
Study on the US Teamsters union cohort ........................ 13
Description of the study and authors’ main results ... 13
Comments on the study ............................................. 13
Study in the unionized US trucking industry ..................... 14
Description of the study and authors’ main results ... 14
Comments on the study ............................................. 14
Study among bus garage workers in Stockholm ............... 17
Description of the study and authors’ main results ... 17

CONTACT Dr. Matthias Mühner Moehner.Matthias@baua.bund.de Federal Institute for Occupational Safety and Health, Division Work and Health, Nöldnerstr. 40-42, 10317 Berlin, Germany
© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
Introduction

In 2012, a working group of the International Agency for Research on Cancer (IARC) upgraded the classification of diesel exhaust (DE) to a Group 1 human carcinogen (Benbrahim-Tallaa et al. 2012; IARC 2014). This decision was mainly based on the findings of the Diesel Exhaust in Miners Study (DEMS). DEMS is the largest epidemiological study on the association between occupational exposure to DE and lung cancer risk based on quantitative exposure assessment. It found a positive association between DE and lung cancer mortality. However, the DEMS primary study results (Attfield et al. 2012; Silverman et al. 2012) are contentious and subject to considerable debate. In particular, the methods involved in the assessment of exposure have been discussed at length. Moreover, the reanalysis of the German Potash Miner cohort study, published shortly after the IARC-decision, could not confirm the DEMS findings (Mohner et al. 2013).

Central to any critical review of occupational health studies, is consideration of the study’s methodological strengths and weaknesses. A first critical issue is the exposure assessment. For the two studies on nonmetal miners, quantitative exposure estimates derived from measurements of respirable elemental carbon (REC) are available. These studies are especially important because of the much higher exposure intensity at underground workplaces in comparison to surface workplaces. In almost all other studies, exposure was assigned based on job title, self-report, or information from next-of-kin. However, DE-exposure is influenced by various factors such as weather conditions, type of vehicle/machine and engine, fuel additives, traffic density, and ventilation. The weaknesses of the exposure assessment in the various DE studies assessing lung cancer risk have already been discussed in depth (Crump 2006; Hesterberg et al. 2006; Gamble 2010; Boffetta 2012a; Crump & Van Landingham 2012; Gamble et al. 2012; Crump et al. 2015, 2016; HEI 2015).

Apart from the assessment of exposure, the selection of study participants, the identification of confounders, and the selection of an appropriate statistical model for the data analysis itself plays an important role for the appraisal of the studies’ findings. We aimed to assess the epidemiological literature on occupational DE exposure and lung cancer with respect to methodological issues and to critically evaluate the evidence for a causal relationship between DE and lung cancer. The focus of this review is primarily on methodological issues other than exposure assessment. Our second aim was to identify studies which might be suitable for quantitative
risk assessment (QRA). Suitability for QRA applies to studies with quantitative exposure assessment, provided that the methodological issues under review support the validity of the study results. Formal QRA generally includes an analysis of the robustness of the final model with regard to alternative approaches for exposure assessment, and is therefore beyond the scope of this review.

Methods

Study selection

The cohort and case-control studies that informed the IARC working group’s decision to upgrade DE into a Group 1 carcinogen were the primary focus of this review (IARC 2014). In addition, we checked the references of other available DE reviews to identify further relevant studies (Stöber & Abel 1996; Bhatia et al. 1998; Lipsett & Campleman 1999; Hesterberg et al. 2006; Gamble 2010; Boffetta 2012a; Gamble et al. 2012; Tsai & Tse 2012; Sun et al. 2014). The most recently published reports on quantitative estimations of the exposure–response relationship were checked in the similar manner (Vermeulen et al. 2014; MacCalman et al. 2015; Neophytou et al. 2016). A complete search for additional publications, which is usually undertaken for a systematic review, was dispensable due to the high actuality in the papers checked.

Case-control studies that lacked information on smoking or did not have an explicit focus on DE exposure were excluded. Three studies that were part of the original IARC report were affected: a study from England/Wales (Coggon et al. 1984), a pooled study of three sub-studies from the United States (Hayes et al. 1989) and a study from Detroit, USA (Swanson et al. 1993). Case-control studies that were nested within cohort studies are discussed in the section on cohort studies of this review. All cohort studies discussed in the IARC report were included in our review, although some studies are not discussed in detail (Menck & Henderson 1976; Bender et al. 1989; Van Den Eeden & Friedman 1993; Boffetta et al. 1988, 2001; Birdsey et al. 2010).

Criteria for the methodological evaluation

There are many instruments for assessing the methodological quality of non-randomized studies (Stroup et al. 2000; Sanderson et al. 2007; Vandenbroucke et al. 2007; Moher et al. 2015). Most of them are simple checklists or scales and, hence, are rather generic in nature. Nevertheless, guidelines for the reporting of meta-analyses of observational studies such as MOOSE, PRISMA, or STROBE describe the main topics to be examined (Stroup et al. 2000; Vandenbroucke et al. 2007; Moher et al. 2015). We adapted these topics to our present review of studies on DE and lung cancer. In view of the comprehensive discussion about weaknesses in the exposure assessment including the development of alternative approaches (Crump 2006; Hesterberg et al. 2006; Gamble 2010; Boffetta 2012a; Crump & Van Ladingham 2012; Gamble et al. 2012; Crump et al. 2015; HEI 2015), we have elected not to discuss this issue in the present review. Any future QRA should proof the robustness of the study results with respect to exposure assessment.

The main questions addressed in this review are the validity of study results and their suitability for a pooled QRA. We therefore examined the following items:

- Were the study participants recruited in an appropriate manner?
- Is the reference group suitable for comparison with the exposed groups in terms of age, birth cohort, and socio-economic status?
- Are there design-specific issues like selection or information bias that could have affected the findings?
- Are there sufficiently detailed and valid data on important confounding factors? Are the statistical methods appropriate to deal with these confounders?
- Might (residual) confounding be an explanation for observed study results?
- Is the statistical modeling of the exposure–response relationship adequate to the data?
- Are the study data suitable to be included into a pooled QRA?

Statistical methods

STATA, release 14 (StatCorp. 2015) was used to recalculate results from specific studies. In particular, we used Poisson regression and conditional as well as unconditional logistic regression. We used Akaike’s Information Criterion (AIC) (Akaike 1974) and the Bayesian Information Criterion (BIC) (Schwarz 1978) for model comparison. Additionally, we implemented two STATA routines, one for the calculation and validation of Standardized Mortality Ratios (SMRs) and their exact confidence intervals based on Poisson distribution (Sasieni 1995) and the other one to perform a meta-analysis (Higgins et al. 2003).

Where available, 95% Confidence Intervals (CI) are provided, unless the respective studies applied a different confidence level (e.g. 90% confidence intervals). As one of the authors (MM) was involved in a large-scale case-control study on DE exposure and lung cancer (Brüseke-Hohlfeld et al. 1997, 1999), the original data from this particular study were used to demonstrate some of the issues discussed in the section on case-control studies.

Results

We identified 18 cohort studies (Table 1) and 13 case-control studies (Tables 10–12) suitable for review. A detailed discussion of the two cohort studies on nonmetal miners, for which quantitative exposure estimates are available, is presented first. These cohorts have the special feature that the exposure intensity at underground workplaces is substantially higher than usually at surface workplaces. We then discuss other cohort studies among railroad workers (3), professional drivers and related jobs (10), dock workers (1), and workers in the construction industry (2). Another six cohort studies mentioned in the IARC report were excluded from detailed review. The reasons for exclusion are shortly explained. Subsequently, we look at hospital-based and population-
<table>
<thead>
<tr>
<th>Reference</th>
<th>Population</th>
<th>Follow-up period</th>
<th>Exposure assessment</th>
<th>Special adjustment factors</th>
<th>Highest exposure category used for comparison</th>
<th>Risk estimate (95% CI)</th>
<th>Considered in IARC's 2012 decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies on non-metal miners</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mühner et al. (2013)</td>
<td>5819 male employees in Potash mining, Germany, employed in 1970 or later</td>
<td>1970–2001</td>
<td>Measurements of TC in 1991; assumption of stable exposure over time</td>
<td>Smoking, Former jobs</td>
<td>Cumulative EC, lagged 5 years, ≥1550 µg/m³-years</td>
<td>OR = 1.04 (0.47–2.27) [nested case-control analysis]</td>
<td>No</td>
</tr>
<tr>
<td>Studies on railroad workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nokso-Koivisto and Puikkala (1994)</td>
<td>8391 members of Finish Locomotive Drivers’ Association, 1953–1991</td>
<td>1953–1991</td>
<td>Job group</td>
<td>≥30 years since first employment as a locomotive driver</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Garshick et al. (2006)</td>
<td>39,388 occupationally active US railroad workers in 1959, with 10–20 years of service, age 40–64 years</td>
<td>1959–1996</td>
<td>Years worked, Years off-work, Smoking, Work in a repair shop</td>
<td>≥20 years worked as engineer or conductor</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Studies on professional drivers and related jobs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steenland et al. (1998)</td>
<td>US Teamsters Union members with at least 20 years of union membership that had died between 1982–1983</td>
<td>Deaths between 1982–1983</td>
<td>EC measurements in trucking industry jobs in 1990; Modelling of historical EC exposure</td>
<td>Smoking, Asbestos</td>
<td>Cumulative EC, lagged 5 years, ≥331 µg/m³-years</td>
<td>OR = 1.64 (1.09–2.49) [nested case-control study based on 994 deaths from lung cancer and 1085 “other deaths”]</td>
<td>Yes</td>
</tr>
<tr>
<td>Garshick et al. (2012)</td>
<td>31,135 male workers employed in US unionized trucking industry in 1985, age ≥40</td>
<td>1985–2000</td>
<td>REC measurements 2001–2006; Modelling of historical REC exposure</td>
<td>Total years of employment</td>
<td>Cumulative REC, lagged 5 years, ≥1803 µg/m³-months</td>
<td>HR = 1.36 (0.98–1.89)</td>
<td>Yes</td>
</tr>
<tr>
<td>Gustafsson et al. (1990)</td>
<td>695 bus garage workers in Stockholm, 1945–1970</td>
<td>1952–1986</td>
<td>Each work period was described in terms of DE exposure intensity (six categories)</td>
<td>–</td>
<td>Highest cumulative DE exposure (DE index &gt;30)</td>
<td>HR = 2.43 (1.32–4.47) [nested case-control analysis]</td>
<td>Yes</td>
</tr>
<tr>
<td>Guberan et al. (1992)</td>
<td>6630 men from Geneva with a professional driver’s licence, 1949–1961</td>
<td>1949–1986</td>
<td>Job group noted on professional driver’s licence</td>
<td>–</td>
<td>≥45 years since start of work as “professional driver”</td>
<td>SMR = 2.59 (1.60–3.96)</td>
<td>Yes</td>
</tr>
<tr>
<td>Rafnsson and Gunnarsdottir (1991)</td>
<td>868 truck drivers in 1939 or later and 726 taxi drivers in 1943 or later from Reykjavik</td>
<td>1951–1988</td>
<td>Job group</td>
<td>–</td>
<td>Truck driver</td>
<td>SMR = 2.14 (1.37–3.18)</td>
<td>Yes</td>
</tr>
<tr>
<td>Guo et al. (2004)</td>
<td>Economically active Fins participating in the National Population Census 1970, including 667,121 men</td>
<td>1971–1995</td>
<td>Application of NO₂-based FINJEM job exposure matrix to longest occupation</td>
<td>Smoking, Asbestos, Quartz dust, SES</td>
<td>Cumulative NO₂, lagged 20 years, ≥10 mg/m³-years (male subgroup)</td>
<td>IRR = 0.95 (0.83–1.10)</td>
<td>Yes</td>
</tr>
<tr>
<td>Soll-Johanning et al. (2003)</td>
<td>18,174 bus drivers and tramway employees, Copenhagen, 1900–1994</td>
<td>1943–1992</td>
<td>Job group</td>
<td>Smoking, Exposure to diesel exhaust in ‘other jobs’</td>
<td>≥20 years employed as driver in public transport, lagged 10 years</td>
<td>OR = 0.54 (0.28–1.03)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(continued)
Table 1. Continued

<table>
<thead>
<tr>
<th>Reference</th>
<th>Population</th>
<th>Follow-up period</th>
<th>Exposure assessment</th>
<th>Special factors</th>
<th>Considered in IARC’s 2012 decision</th>
<th>Risk estimate</th>
<th>Special adjustment factors</th>
<th>Risk estimate (95% CI)</th>
<th>Considered in IARC’s 2012 decision</th>
<th>Risk estimate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petersen et al. (2010)</td>
<td>2037 male urban bus drivers, Denmark, 1978</td>
<td>1979–2003</td>
<td>Job group</td>
<td>Smoking</td>
<td>Yes</td>
<td>HR = 0.8 (0.5–1.4)</td>
<td>No</td>
<td>Yes</td>
<td>HR = 0.8 (0.5–1.4)</td>
<td></td>
</tr>
<tr>
<td>Hansen (1993)</td>
<td>14,225 truck drivers and 43,024 other unskilled laborers</td>
<td>1970–1980</td>
<td>Job group</td>
<td>Yes</td>
<td>Yes</td>
<td>SMR = 1.60 (1.26–2.00)</td>
<td>No</td>
<td>Yes</td>
<td>SMR = 1.60 (1.26–2.00)</td>
<td></td>
</tr>
<tr>
<td>Wong et al. (1998)</td>
<td>34,156 male members of US heavy construction equipment operators union</td>
<td>1964–1978</td>
<td>Job group</td>
<td>Yes</td>
<td>Yes</td>
<td>SMR = 1.07 (0.91–1.25)</td>
<td>No</td>
<td>Yes</td>
<td>SMR = 1.07 (0.91–1.25)</td>
<td></td>
</tr>
</tbody>
</table>

**Diesel Exhaust in Miners Study**

**Description of the study and authors’ main results**

The DEMS is the largest epidemiological study using quantitative exposure estimates to evaluate occupational DE exposure and lung cancer risk. This retrospective cohort mortality study involved 12,315 workers at eight US nonmetal mining facilities. Data were analyzed using a cohort, as well as a nested case-control approach (Attfield et al. 2012; Silverman et al. 2012). Methodologically, DEMS offers advantages over most other studies on DE and lung cancer in that it used DE exposure measurements to construct a job-exposure matrix (JEM). The JEM is based on measurement surveys of exposure to respirable elemental carbon (REC), carried out between 1998 and 2001. REC exposure levels prior to these surveys were estimated; the REC measurements were weighted according to the estimated relative trend in carbon monoxide (CO) exposure, which was determined using 1976–2001 CO measurement data and 1947–2001 information on diesel equipment and mine ventilation. REC exposure intensity differed by nearly two orders of magnitude between underground and surface workers (Coble et al. 2010; Stewart et al. 2010, 2012; Vermeulen et al. 2010a, 2010b). Results showed that in comparison to the general population, miners had an increased risk of lung cancer (overall SMR for lung cancer 1.26). However, despite their much higher DE exposure intensity, the lung cancer risk for ever-underground workers was somewhat lower than that of surface-only workers (SMR 1.21 vs. 1.33) (Attfield et al. 2012). Even the primary (a priori defined) internal cohort analysis did not show an association between cumulative REC exposure or average REC intensity and lung cancer mortality (Attfield et al. 2012). A positive exposure–response relationship was only seen when the (time-dependent) binary variable “work location” was included in the model. Subsequently, the investigators also incorporated this variable in their case-control approach.

In the case-control analysis, a statistically significant positive association between lung cancer and increasing cumulative REC exposure, as well as average REC intensity, was observed (Silverman et al. 2012). The OR comparing the highest and lowest quartile of cumulative REC was 2.83 (95% CI 1.28–6.26), adjusted for work location cross-classified by smoking status and smoking intensity. The DEMS authors concluded that their study findings provided evidence that exposure to DE increases the risk of mortality from lung cancer.

The conflicting results of different methodological approaches applied in DEMS led to several critical commentaries and letters to the editor (Boffetta 2012b; Hesterberg et al. 2012; McClellan 2012; Morfeld 2012; Möhner et al. 2012; Tse & Yu 2012; Pallapies et al. 2013). Following this debate, an expert panel was set up by the Health Effects Institute (HEI) to evaluate the DEMS results (HEI 2015). Additional analyses of the data were performed to verify the robustness of the risk estimates with respect to alternative approaches for
exposure assessment (Crump et al. 2015, 2016) and time-related factors (Moolgavkar et al. 2015).

Comments on the study
External cohort analysis. DEMS hypothesized that a positive exposure–response relationship exists between DE and lung cancer. However, contradicting this is the study finding of a similar increase in the standardized mortality ratio (SMR) for lung cancer in surface-only and ever-underground workers despite differing exposure intensities. Exposure intensity was 75 times higher for ever-underground workers compared with surface-only workers, whereas the SMR was slightly lower in ever-underground workers than in surface-only workers (1.21 vs. 1.33) (Attfield et al. 2012). The HEI panel noted in its recent report that "these results should not be over-interpreted" because “these analyses cannot take into account any covariates (e.g., smoking, other occupational exposures)” (HEI 2015). However, the distribution of controls in the case-control study (Table 2 in Silverman et al. 2012) shows that there were significantly more never-smokers (34% vs. 22%) and significantly less ever smokers (6% vs. 14%) among surface-only workers than among ever-underground workers. Thus, the increased SMR in surface-only as compared to ever-underground workers, in a situation of heavily diverging DE exposure intensities, is probably not attributable to smoking. Of note, mortality due to pneumoconiosis was considerably higher in ever-underground workers (SMR = 16.21 vs. 6.13), indicating that underground workers experienced increased disease risk related to dust exposure as compared to surface workers.

Internal cohort analysis. The primary (a priori defined) internal cohort analysis did not show an association between cumulative REC exposure or average REC intensity and lung cancer mortality. The risk estimates for the upper three quartiles of cumulative REC exposure, lagged 15 years, were all less than 1.0, i.e., HR = 0.58, 0.71, and 0.93 (Attfield et al. 2012). Such a gradient is typical for a healthy-worker effect (HWE) if one considers the surface-only workers as the unexposed group. The drop of risk in the second quartile might be related to a change of a certain part of workers complying the health requirements for underground work from surface to underground (Möhner 2016b). The DEMS authors observed an increasing lung cancer risk primarily in the ever-underground workers in relation to increasing cumulative REC. But, this observation might also be triggered by a HWE, because the workers in the reference group have spent less than one year at underground workplaces (cumulative REC exposure <108 µg/m²-years).

Including the time-dependent binary work location in the Cox model, a positive exposure–response relationship was observed. The OR in the highest exposure quartile, lagged 15 years, as compared to the lowest quartile changed to 1.39 (95% CI 0.78–2.48). However, the variable "work location" contains the bulk of the exposure information and, hence, adjusting for this variable would lead to overadjustment bias (Schisterman et al. 2009).

Nested case-control analysis. The simultaneous inclusion of work location and cumulative exposure (or average intensity) is also an issue in the case-control analysis. To verify the selected statistical model, we performed a comparison with the other possible models listed in Table 2. Without admission to the original database, this analysis could only be carried out using unconditional logistic regression based on the published frequencies of cases and controls (Table 2 in Silverman et al. 2012). The unconditional approach leads to slightly more conservative estimates in comparison to the conditional approach (Breslow & Day 1980). However, the matching ratio in DEMS was 1:4 and, consequently any deviation should be small. We calculated Akaike’s information criterion (AIC) and the Bayesian information criterion (BIC) to account for both the estimation error and the model bias. As shown by Greenland and colleagues, conditional logistic regression can exhibit considerable bias when certain types of matched sets are infrequent or when the model contains too many parameters (Greenland et al. 2000). The DEMS is, strictly speaking, a pooled study of eight cohorts that differ considerably by year of dieselization, by range of exposure, by size of the workforce, and possibly also by other factors like history of former employment in other mines. The analysis of Moolgavkar and colleagues pointed out that the SMRs for lung cancer, as well as the slopes for the exposure–response relationship, are different between mine types (Moolgavkar et al. 2015). Therefore, together with year of birth (divided into 5-year birth cohorts) nearly 20 additional parameters need to be estimated in a corresponding unmatched study design, i.e. altogether more than 40 parameters. Hence, the corresponding estimates are imprecise. The same problem exists for a conditional logistic regression model under the matched design, referred to as “sparse-data” bias (Greenland et al. 2000). Therefore, the BIC is an appropriate criterion to select a parsimonious model, because it is

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables and interactions</th>
<th>d.f.</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
<td>–</td>
<td>1</td>
<td>932.126</td>
<td>935.4144</td>
</tr>
<tr>
<td>M1</td>
<td>Work location</td>
<td>2</td>
<td>934.088</td>
<td>940.6648</td>
</tr>
<tr>
<td>M2</td>
<td>Work location + smoking status</td>
<td>5</td>
<td>883.7319</td>
<td>900.1732</td>
</tr>
<tr>
<td>M3</td>
<td>Work location + smoking intensity</td>
<td>6</td>
<td>890.4929</td>
<td>910.2225</td>
</tr>
<tr>
<td>M4</td>
<td>Work location + smoking status + smoking intensity</td>
<td>7</td>
<td>881.3167</td>
<td>904.3345</td>
</tr>
<tr>
<td>M5</td>
<td>Work location + (smoking status × smoking intensity)</td>
<td>9</td>
<td>885.1182</td>
<td>914.7126</td>
</tr>
<tr>
<td>M6</td>
<td>Work location × smoking status × smoking intensity</td>
<td>16</td>
<td>884.8769</td>
<td>937.4892</td>
</tr>
</tbody>
</table>

*Number of cases and controls are taken from Table 2 in Silverman et al. (2012).
AIC: Akaike’s information criterion; BIC: Bayesian information criterion; d.f.: Degrees of freedom; +: Only main effects of the variables are included into the model; ×: Main effect and interaction terms are included into the model.
Values for the best fitting model are given in bold.
more sensitive with respect to the inclusion of further parameters in comparison to AIC. We calculated BIC as:

$$\text{BIC} = -2 \ln(\text{likelihood}) + \ln(N)k$$

where $k$ is the number of parameters in the model and $N$ is the number of matched sets (i.e., the number of cases under an unconditional approach). The calculation of AIC and BIC is a standard tool in many statistical software packages, as for example STATA.

The reported numbers of cases and controls (Table 2 in Silverman et al. 2012) enabled us to apply unconditional logistic regression models based on the variables work location, smoking status and smoking intensity. Among different reasonable models, we aimed to determine the best model with respect to AIC and BIC. Our simplest model (M2) incorporates work location and smoking status as independent variables. Model M6 incorporates the variable work location cross-classified by smoking status and intensity. Hence, model M6 is most comparable to the conditional logistic regression model originally used to analyze the DEMS data (i.e. Silverman model in Table 3). Beside the switch from the conditional to the unconditional approach, the original model differs from model M6 in the addition of the quartiles of cumulative REC exposure and two other confounding factors (history of respiratory diseases and history of other high-risk occupations).

The models M2 (work location and smoking status) and M4 (work location, smoking status and smoking intensity) both yield lower values for AIC and BIC in comparison with model M6 (work location cross-classified with smoking status and smoking intensity) (Table 2). Therefore, both models are better than model M6 in the sense of the two information criteria and they reduce the number of model parameters considerably (reduction by 11 or 9 parameters, respectively).

We calculated the predicted lung cancer risks based on the alternative models and compared them to the original estimates derived by the DEMS investigators from models adjusted for cumulative DE exposure (Table 2 in Silverman et al. 2012) (Table 3). For surface-only workers, Table 3 shows only marginal differences between the ORs based on model M6 and the original DEMS model. This observation suggests that the additional adjustment for history of respiratory diseases and history of high-risk job in the original DEMS analysis has only little impact on risk estimates and it underlines that the estimates derived by unconditional and conditional logistic regression are very similar. In contrast, the two sets of estimates differ by a factor of about two for ever-underground workers. The estimates of the original model are lower than the estimates of model M6 that excludes the categorical REC variable. However, the vast majority of ever-underground workers fall into the two upper REC exposure quartiles, whereas almost all surface-only workers are classified into the two lower quartiles. Despite that, the cumulative REC exposure was treated as an independent variable in the approach of the DEMS authors. We could thus easily derive risk estimates comparable to that for model M6 by multiplying the ORs for smoking, cross-classified with work location (Table 2 in Silverman et al. 2012) with the ORs for the quartiles of cumulative REC (Table 3 in Silverman et al. 2012). The multiplier is approximately one for surface-only workers (ORs of 0.74 and 1.00 for lower REC quartiles) and approximately two for ever-underground workers (ORs of 1.54 and 2.83 for upper REC quartiles). Thus, the difference between model M6 and the original DEMS model is due to the additional adjustment for cumulative REC exposure in the DEMS model.

It should be noted that the mean HR for the time-dependent variable work location in the internal cohort analysis was of similar magnitude (1.9, range: 1.64–2.28) (Attfield et al. 2012). We can conclude from this, that the simultaneous modeling of the highly correlated variables work location (cross-classified by smoking) and cumulative REC leads to overadjustment bias. Our calculations have shown that the estimates for the interaction term (between smoking and work location) in the Silverman model and the estimates of the (main) effect of high level of REC exposure compensate each other. Thus, the effect of REC in the Silverman model is an apparent effect only.

According to AIC and BIC, the models M2 and M4 are best suited for the estimation of lung cancer risk by work location. These models simply adjust for smoking habits. In the same manner, the models estimating lung cancer risk in relation to cumulative REC should not include the variable work location.

We repeated our approach based on the distribution of cases and controls by tertiles of cumulative REC exposure and smoking intensity (Table 6 in Silverman et al. 2012). The DEMS-authors reported a $p$-value of 0.086 for interaction between smoking intensity and cumulative REC, lagged 15 years. Our unconditional approach yields a $p$-value of 0.089 for these interaction terms (which again underlines the similarity of both approaches for the given study data). The comparison of the models revealed that the model without the interaction term (M8) is superior to the full model

---

**Table 3. Comparison of odds ratios from different models for lung cancer risk in DEMS by work location and smoking status/intensity.**

<table>
<thead>
<tr>
<th>Smoking status/intensity</th>
<th>Surface-only worker</th>
<th>Ever-underground worker</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (M2)$^b$</td>
<td>OR (M4)$^b$</td>
</tr>
<tr>
<td>Never smoker</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Former (0, 1)</td>
<td>3.47</td>
<td>2.79</td>
</tr>
<tr>
<td>Former (1, 2)</td>
<td>3.47</td>
<td>3.40</td>
</tr>
<tr>
<td>Former (2, on)</td>
<td>3.47</td>
<td>5.46</td>
</tr>
<tr>
<td>Current (0, 1)</td>
<td>6.91</td>
<td>5.26</td>
</tr>
<tr>
<td>Current (1, 2)</td>
<td>6.91</td>
<td>6.42</td>
</tr>
<tr>
<td>Current (2, on)</td>
<td>6.91</td>
<td>10.30</td>
</tr>
<tr>
<td>Unknown</td>
<td>3.63</td>
<td>3.63</td>
</tr>
</tbody>
</table>

$^a$Odds ratios from Table 2 in Silverman et al. (2012).

$^b$Models as described in Table 2.
and surface workplaces with respect to REC intensity, the average REC intensity primarily reflects the share of surface work in the overall exposure duration. The interpretation of average DE exposure intensity in DEMS thus differs from that of usual average values. Moreover, 22% of ever-underground workers started their work at surface and switched to underground later. Hence, low values of average REC intensity reflect a healthy-worker effect among these surface-first workers (Möhner 2016b).

Confounding by other factors. In their nested case-control study, the DEMS investigators ascertained data on several potential confounders. However, their final model included only two confounders that changed the ORs for DE metrics by more than 10% (history of respiratory disease ≥5 years before index date and history of a high-risk job for lung cancer for ≥10 years). In contrast, Crump and colleagues showed that an additional adjustment for radon exposure reduces the estimates for the DE effect markedly when some alternative metrics for REC were applied (Crump et al. 2015). In our view, the selection of confounding factors should also be based on appropriate model selection criteria. Beside the information criteria mentioned above, cross validation would have been an appropriate approach for model selection (Arlot & Celisse 2010).

Exposure Lagging. The DEMS finding should also be interpreted with an understanding of the exposure lagging. There are two reasons for implementing exposure lagging into a study; to make sure an adequate latency period is

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables and interactions</th>
<th>d.f.</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
<td>Tertiles of cumulative REC</td>
<td>1</td>
<td>932.1262</td>
<td>935.4144</td>
</tr>
<tr>
<td>M7</td>
<td>Tertiles of cumulative REC</td>
<td>3</td>
<td>932.9279</td>
<td>942.7927</td>
</tr>
<tr>
<td>M8</td>
<td>Tertiles of cumulative REC + smoking intensity</td>
<td>7</td>
<td>890.2927</td>
<td>913.3106</td>
</tr>
<tr>
<td>M9</td>
<td>Tertiles of cumulative REC x smoking intensity</td>
<td>15</td>
<td>892.5585</td>
<td>941.8825</td>
</tr>
</tbody>
</table>

*Number of cases and controls are taken from Table 6 in Silverman et al. (2012). AIC: Akaike’s information criterion; BIC: Bayesian information criterion; d.f.: Degrees of freedom.

Values for the best fitting model are given in bold.

**Table 5.** Comparison of odds ratios from different models for lung cancer risk in DEMS by work location and smoking intensity.

<table>
<thead>
<tr>
<th>Smoking packs/day</th>
<th>Tertile 1, 0 to &lt;8 μg/m³-years</th>
<th>Tertile 2, 8 to &lt;304 μg/m³-years</th>
<th>Tertile 3, ≥304 μg/m³-years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (M8)b</td>
<td>OR (M9)b</td>
<td>OR (Silverman)a</td>
</tr>
<tr>
<td>[0]</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(0, 1)</td>
<td>3.40</td>
<td>4.80</td>
<td>6.25</td>
</tr>
<tr>
<td>[1, 2)</td>
<td>4.92</td>
<td>7.31</td>
<td>10.16</td>
</tr>
<tr>
<td>(2, ∞)</td>
<td>7.65</td>
<td>16.98</td>
<td>26.79</td>
</tr>
<tr>
<td>Unknown</td>
<td>3.60</td>
<td>3.15</td>
<td>4.13</td>
</tr>
</tbody>
</table>

*Odds ratio from Table 6 in Silverman et al. (2012). Models as described in Table 4.

**Table 6.** Lung cancer risk estimates for various categorizations of cumulative REC exposure in the DEMS.

<table>
<thead>
<tr>
<th>Exposure/reference group</th>
<th>Unadjusted for smoking</th>
<th>Adjusted for smoking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Reference: surface-only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever-underground</td>
<td>1.03 0.74–1.43</td>
<td>0.94a 0.67–1.32</td>
</tr>
<tr>
<td>Reference: Cumulative REC, lagged 15 years</td>
<td>0.89 0.60–1.31</td>
<td>0.93b 0.62–1.38</td>
</tr>
<tr>
<td>Tertile 2 (8–304 μg/m³-years)</td>
<td>1.26 0.85–1.86</td>
<td>1.27b 0.85–1.89</td>
</tr>
<tr>
<td>Tertile 3 (≥304 μg/m³-years)</td>
<td>0.72 0.46–1.12</td>
<td></td>
</tr>
<tr>
<td>Reference: Cumulative REC, unlagged, quartile 1 (0–19 μg/m³-years)</td>
<td>1.03 0.65–1.62</td>
<td></td>
</tr>
<tr>
<td>Quartile 2 (19–246 μg/m³-years)</td>
<td>1.00 0.64–1.57</td>
<td></td>
</tr>
<tr>
<td>Quartile 3 (246–964 μg/m³-years)</td>
<td>0.71 0.45–1.10</td>
<td></td>
</tr>
<tr>
<td>Quartile 4 (≥964 μg/m³-years)</td>
<td>1.01 0.64–1.58</td>
<td></td>
</tr>
<tr>
<td>Reference: Cumulative REC, lagged 15 years, quartile 1 (0–3 μg/m³-years)</td>
<td>1.31 0.83–2.07</td>
<td></td>
</tr>
<tr>
<td>Quartile 2 (3–72 μg/m³-years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quartile 3 (72–536 μg/m³-years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quartile 4 (≥536 μg/m³-years)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aAdjusted for smoking status and smoking intensity.

*bAdjusted for smoking intensity.

Derived by unconditional logistic regression, number of cases and controls are taken from Tables 2, 3, and 6 in Silverman et al. (2012).
incorporated into the analysis and to prevent a HWSE. The DEMS-authors published results for exposure unlagged and lagged 15 years. A lag of 15 years was selected as it yielded the best fit to the data in comparison with other lag times tested. However, for lung cancer the interval between diagnosis and death is small, the median is about 10 months (Compton et al. 2012). A lag-time of 5 years would be sufficient to take into account premature termination of exposure (Checkoway et al. 1990). Furthermore, a lag-time of 15 years seems too long with respect to the latency period. It is well known that smoking cessation reduces lung cancer risk in comparison to continuing smoking after just a few years (Peto et al. 2000; Fry et al. 2013). Hence, one cannot rule out that REC exposure that occurred in the recent 15 years (except the five most recent years) is of importance and it should thus not be easily disregarded.

**Summary.** At present, the DEMS data represent the most important basis for the quantitative analysis of the relationship between exposure to DE and lung cancer. The advantage of this study is its wide range of exposure intensity at different workplaces in the mines, the availability of quantitative exposure estimates, and the availability of information on previous employments and on other potential confounders like smoking. In view of the huge difference between surface and underground workplaces with regard to DE exposure intensity, surface-only workers are the natural reference group for the investigation of dose–response relationships. Selection bias seems not to play any role. However, the major flaw of the study is the inadequate model choice, resulting in a strong overadjustment bias combined with sparse-data bias. At present, the DEMS does not add evidence to an exposure–response relationship between DE and lung cancer. A reanalysis of the original data is recommended.

**German Potash Miners Cohort Study**

**Description of the study and authors’ main results**

Potash has been extracted in underground mines in the South Harz Mountains area of Germany for several decades. Dieselization of the mining technology was completed in 1969. From that time to the mine closure in 1991, nearly 6000 miners were exposed to DE.

Exposure assessment was based on measurements of the concentration of total carbon (TC) in the airborne fine dust fraction by coulometric analysis, carried out in 1991 (Dahmann et al. 1996). To a lesser extent, also measurements of elemental carbon (EC) were available. As the mining technology and the mining equipment remained fairly stable since 1969, measurements from 1991 have been used to design a job-exposure-matrix for the complete study time.

Follow-up mortality data on the cohort covers the years from 1970 to 2001 (Sävéri et al. 1999; Neumeyer-Gromen et al. 2009). In sum, the study recorded 61 lung cancer cases, resulting in a significantly reduced SMR [SMR = 0.73 (95% CI: 0.57–0.93)]. Internal analysis showed an increased risk of lung cancer in the fifth quintile of cumulative exposure to total carbon [HR = 2.28 (95% CI: 0.87–5.97)]. Adjustment for time-since-hire increased the risk estimates further, although only the binary exposure variable results were initially reported [increase of HR from 1.28 up to 2.53 (95% CI: 1.13–5.69)]. Later, it was suspected that some miners with an occupational history of uranium mining could have influenced the results. A reanalysis was subsequently performed which took previous uranium and coal mining into consideration (Möhner et al. 2013). This analysis did not show any notable association between cumulative exposure to REC and lung cancer risk [RR = 1.06 (95% CI: 0.50–2.23) for the comparison of the highest quartile of REC exposure with the lowest quartile, adjusted for smoking and previous employment]. Former employment in uranium mining did increase lung cancer risk significantly, but it was not correlated with the DE exposure (due to the fact that the former uranium miners switched to the potash mine before dieselization). Hence, the adjustment for former employment in uranium mining did not change the DE-related risk estimate.

**Comments on the study**

The conflicting results of the two analyses were the reason for the exclusion of this study from a meta-regression of cohort data to estimate a dose–response relationship for DE (Vermeulen et al. 2014). Therefore, this comment provides an explanation for the differing results in the separate analyses of the same study data.

In the original analysis, the risk estimate for the dichotomized DE exposure variable increased with the inclusion of time since first exposure (TSFE) in the Cox regression model that used attained age as the time-scale (Neumeyer-Gromen et al. 2009). However, TSFE is strongly correlated with the cumulative exposure variable. This fact is also reflected in the corresponding estimate for the impact of TSFE. The estimated hazard ratio for TSFE was HR = 0.79 (95% CI: 0.76–0.82) per year in the specified model (own recalculations of the study data). Taking into account that the chosen single cut-point for the cumulative exposure of 4.9 mg/m³-years total carbon (corresponding to 3.1 mg/m³-years REC; 1 mg/m³-year is used as measure for cumulative exposure. It equates to a time-weighted average of 1 mg/m³ over one year of work, i.e. 220 shifts of 8 h each) corresponds to about 20 years of exposure in the production area, it is clear that this estimation procedure leads to incorrect results.

Instead of a single cut-point, the authors of the original analysis also considered quintiles of the cumulative exposure variable (Table 5 in Neumeyer-Gromen et al. 2009). However, results were only reported for a Cox model with TSFE as the time scale. The common recommendation is to use the time-scale that is biologically most relevant (Korn et al. 1997; Thiebaut & Benichou 2004; Griffin et al. 2012). Thus, attained age instead of TSFE should have been used as the time scale. The specified model included age at study entry as a continuous adjustment variable. The age range of the cohort at dieselization was from 14 to 65 years. The corresponding risk estimate was HR = 1.13 (95% CI: 1.09–1.17) per year of age (own recalculations of the study data). However, the fit of a model with a linear age effect in this broad age range cannot adequately take into account the known strong non-linear.
effect of age on lung cancer risk and is therefore inaccurate. Not least because of these methodological shortcomings, a reanalysis of the study data was necessary.

In the reanalysis of the data, age was used as the time scale (Möhner et al. 2013). It demonstrated that former employment in uranium mines significantly increased the risk of lung cancer. However, in contrast to what was expected, this status (i.e. ever versus never engaged in uranium mining) did not modify the risk estimates for DE. In other words, former uranium mining and DE were acting as independent risk factors. The reanalysis did not corroborate the increase in lung cancer risk with high cumulative DE exposure in the original analysis. Both a cohort approach and a nested case-control approach were used for data analysis, and neither approach showed increase in lung cancer risk. Additionally, using the case-control approach, a comparison between high-DE exposed versus low-DE exposed miners was performed. The cutoff point was shifted upwards (until the number of cases was not less than five) in order to determine a potential threshold value. The best fit with respect to AIC and BIC was derived for the cutoff point 2.9 mg/m³-years (OR = 1.81; 95% CI: 0.57–4.52). However, the difference to other models in terms of AIC was negligible (max [AIC(x) – AIC(2.9)] < 0.8, for x in the range of 2.0–3.0). It was then concluded in a conservative manner that below 2.5 mg/m³-years of cumulative REC exposure, there should be no elevated lung cancer risk.

Contrary to the hypothesis of Vermeulen et al. (2014), the narrower range of REC intensity in the Potash Miners Cohort, as compared to DEMS, should not mask a causal association between DE and lung cancer, if it were present. The reanalysis of the study data rather suggests that an association between DE and lung cancer is not present.

Summary. The data of the German Potash Miners Cohort are suitable for QRA with respect to DE and lung cancer, despite the smaller number of lung cancer cases in comparison to DEMS. The underground workplaces in Potash mines examined in both studies have similar DE exposure intensities. However, the range of cumulative exposure in the Potash study is not as wide as almost all workplaces were located underground. Information on confounding factors in the Potash study were restricted to entry into the mine, data on former mining and crude information on smoking status. Nevertheless, the results of the reanalysis of this cohort support the notion that a clear relationship between DE and lung cancer is absent, at least in the range of a cumulative REC exposure up to 2.5 mg/m³-years. The formerly suspected strong relationship between DE and lung cancer based on the original analysis of the long-term follow-up of the cohort is misleading due to methodological shortcomings such as adjustment for time since hire.

Cohort study in Canadian retired railway workers

Description of the study and authors’ main results

Analysis is based on a cohort of 43,826 male pensioners of the Canadian National Railway Company (Howe et al. 1983). A total of 17,838 pensioners died between 1965 and 1977. Information on smoking status was not available. Jobs were classified by the level of exposure to DE and coal dust. Job information was based on the last known job. The SMR for all causes of death was 0.95 (95% CI: 0.93–0.96; own calculation based on information in study report). The lung cancer risk was slightly elevated [SMR = 1.06 (95% CI: 0.99–1.13)]. A comparison of job groups yielded a significantly increased relative risk for a probable exposure to DE as well as to coal dust (RR of 1.35 in both analyses).

Comments on the study

The identical increase in lung cancer risk with probable exposure to DE and coal dust should draw our attention also to further factors related to exposed jobs. A useful classification might be to distinguish between mobile workers (i.e. train crews) and stationary workers. Mobile workers are often on the rail for several days. Their lifestyle might be quite different from that of stationary workers, especially with respect to smoking and dietary habits. This hypothesis is supported by evidence of elevated risks for rectum cancer in engineers [SMR = 1.70 (95% CI: 1.07–2.57)] and colorectal cancer in motorman [SMR = 3.15 (95% CI: 1.63–5.51)] (own computations based on information in study report). Moreover, as shown in a case-control study among current and retired US railroad employees, asbestos exposure plays an important role in the mortality of railroad workers (Schenker et al. 1986) and should thus be considered as a confounding factor in analyses on lung cancer in railroad workers.

Summary. Lung cancer risk is only slightly elevated in the cohort in comparison with the general population. The reference group for the internal comparison seems to be different with respect to lifestyle factors, but information on important confounding factors is not available. Hence, the results of the study are of restricted use for the clarification of the relationship between DE and lung cancer. Due to missing quantitative measures of exposure, the study is not useable for QRA.

Cohort study among Finnish locomotive drivers

Description of the study and authors’ main results

This was a retrospective cohort study where the cancer incidence of 8391 members of the Finnish Locomotive Drivers’ Association was analyzed for the period from 1953 to 1991 (Nokso-Koivisto & Pukkala 1994). The most notable result was the significantly elevated risk for mesothelioma [SMR = 4.05 (95% CI: 1.75–7.97)]. In contrast, lung cancer incidence was not elevated [SMR = 0.86 (95% CI: 0.75–0.97)] and lung cancer risk increased only slightly with calendar time up to 0.93 (95% CI: 0.77–1.10) in the period 1979–1991. Analysis by the duration of follow-up revealed that 6 cases occurred up to 14 years after first employment or 1953, whichever was later [SMR = 1.02 (95% CI: 0.37–2.22)].

Comments on the study

Diesel locomotives were introduced in the Finnish railway industry in the 1950s. Hence, the vast majority of the cohort
members were exposed to DE. Taking into account the mean age of onset of lung cancer is about 65, many of the subjects with lung cancer were likely to have also worked on a steam locomotive prior to dieselization. The study authors pointed out in their methods section, that “all the Finnish locomotive drivers start their training at about the same age, from 16 to 20”. Hence, the 6 cases, which occurred up to 14 years since entry into the study (1953–1966), occurred very likely among those workers with a preceding exposure in the steam locomotive era. Moreover, the authors suspect “that the increased smoking in the less draughty cabins of diesel locomotives with less physical work and more passive monitoring would have increased the risk of lung cancer close to the national average”, which could explain the slight increase of SIR by calendar time. Furthermore, they refer to a cross-sectional study of Finnish locomotive drivers in 1976, which showed a similar distribution of smoking status with respect to the general population.

Summary. For this incidence-based cohort study of engineers, individual information on occupational exposures and confounding factors are not available. Hence, the study can only offer a rough impression about lung cancer risk in comparison to the general population. The overall SIR for lung cancer is significantly below one and a trend with follow-up time was not observed. Hence, the study does not support the hypothesis of a causal relationship between DE exposure and lung cancer risk. Due to missing quantitative measurements of exposure, it is not useable for QRA.

Cohort study among US railroad workers

Description of the study and authors’ main results
This study included ~55,000 US railroad workers followed-up from 1959 to 1996 (Garshick et al. 2004, 2006; Laden et al. 2006). The study is based on several predecessor studies: a pilot study (Schenker et al. 1984), a first follow-up until 1980 (Garshick et al. 1988), a nested case-control study ascertaining information on smoking (Garshick et al. 1987a), studies estimating the DE exposure in railroad workers (Woskie et al. 1988a, 1988b; Schenker et al. 1990), a reanalysis of the first follow-up (Crump 1999), additional analyses by the Diesel Epidemiology Expert Panel convened by the Health Effects Institute (HEI 1999), and a methodological paper dealing with the smoking adjustment in the cohort based on the data gathered in the case-control study (Larkin et al. 2000).

The cohort represents a sample of male workers aged 40–64 years with 10–20 years of railroad service, who were employed in one of 39 pre-specified jobs at start of follow-up in 1959, when dieselization of locomotives in the US railroad industry (started after World War II) was almost complete. Inclusion sampling into the study cohort was weighted according to job group; every third engineer (engineers and firemen), every third conductor (conductors, brakemen, and hostlers), all repair shop workers (shop supervisors, machinists and electricians), and a reference group of all ticket agents, station agents, signal maintainers and every fourth clerk with comparatively low DE exposure. Exposure assessment was based on the classification of job groups into “DE-exposed” (“engineers”, “conductor”, and “shop workers”) and “unexposed”, using a measurement survey on job group-specific concentrations of respirable particles (Woskie et al. 1988a, 1988b). Shop workers with potential for past heavy asbestos exposure before 1959 due to repair of steam-powered locomotives (e.g. boilermakers and helpers) and other workers potentially exposed with asbestos while repairing passenger cars, steam pipes or constructing and repairing railroad buildings were excluded from the cohort. Overall, approximately 75% of the workers were in DE exposed jobs and 25% were in low- or non-exposed jobs.

Adjusting for attained age, years of employment in the railroad industry, and time since last worked as time-dependent covariates, the investigators calculated a relative risk of lung cancer mortality of 1.40 (95% CI: 1.30–1.51) for engineers or conductors that worked on trains with exposure to DE (Garshick et al. 2004). However, no exposure–response relationship was found in exposed workers, i.e. lung cancer mortality did not increase with increasing years of work in these jobs. Furthermore, shop work for which DE exposure was also high was not associated with lung cancer. Stratified analysis by age at study entry showed that lung cancer risk was increased preferentially in younger engineers and conductors. This finding was interpreted by the investigators as evidence of longer exposure to DE being important for lung cancer development. After imputation of smoking data, the investigators stated that the relative lung cancer risk for DE exposed railroad workers (when applying an exposure lag of 5 years) decreased from 1.35 (95% CI 1.24–1.46) to 1.22 (95% CI: 1.12–1.32) with adjustment for smoking (Garshick et al. 2006). Additionally, the investigators refined the “years of work in DE exposed job” further, using historical information on the use of diesel locomotives since 1945 and information on emissions from diesel-operated locomotives (Laden et al. 2006). Using the new “weighted years of work in DE job”, the investigators observed that the increase of lung cancer risk in relation to a DE exposed job was stronger for workers hired from 1945 to 1949 than for workers hired from 1939 to 1944. The authors also described an exposure–response relationship in the group of workers that was hired after 1945.

Comments on the study

Adjustment for years worked. For the analysis based on the follow-up period until 1996, the investigators used proportional hazard models to estimate relative risks for various DE-exposed job groups and adjustment for HWSE was considered. In their report, the investigators write, “To account for a healthy worker survivor effect, we included time-varying variables for total years worked and for years off work (usually time after retirement) in survival models” (Garshick et al. 2004). However, it is questionable if the adjustment for years worked really is an adjustment for healthy worker survivor bias. A healthy worker survivor bias may occur in occupational studies due to the tendency for unhealthy individuals to leave work earlier, and consequently accrue less exposure, compared with their healthier counterparts (Buckley et al. 2015). The problem in this study is rather left truncation due to the selection of prevalent hires into the cohort...
(Applebaum et al. 2011). Ideally, in occupational epidemiology, incident hires form the basis of a study cohort, whereas the US railroad worker cohort included prevalent hires of different birth cohorts.

Selection bias seems to be strongest for the oldest workers, as suggested by the highest median age at retirement in the oldest birth cohort. Yet, by study design (year of hire for all subjects between 1939 and 1949), the oldest birth cohort accrued the least years of service in the railroad companies, while the youngest workers accrued the longest service. As reported by the authors, lung cancer mortality was inversely related to total years worked (RR = 0.97 for each additional year of employment). This result suggests that the youngest birth cohort was particularly healthy, although due to left truncation, this actually held true for the oldest birth cohort. Total years worked was thus not an appropriate variable to control for HWSE, at least under the given study design. Secondly, this adjustment was inappropriate as only workers with 10–20 years of service entered the study in the year 1959, when exposure to DE also started. Hence, all birth cohorts were homogeneous with respect to their pre-diesel experience and the difference in total years of work between birth cohorts was highly correlated with the cumulative exposure. Consequently, adjusting for total years worked results in further overadjustment bias. This bias could have been avoided if only years of employment in the pre-diesel era were included in the model as a confounder.

The situation for adjustment for years off work was similar. The mean age at entry into the study for the oldest birth cohort was virtually the same as the retirement age of the youngest birth cohort. As the median age at retirement was higher in the oldest birth cohort than in younger birth cohorts, adjustment for years off work is not justified due to the strong correlation between birth cohort and cumulative exposure.

**Comparability between mobile and stationary workers.**

Crum’s reanalysis of the cohort study (Crum 1999) revealed an increased lung cancer risk for exposed train crews (i.e., engineers, firemen, conductors and brakemen) in comparison to non-exposed stationary workers (i.e., station clerks and signalmen) and exposed stationary shop workers. This observation was later confirmed by Garshick et al. (2004). Further to that, Crump also reported a higher mortality of the group of mobile workers for heart disease and liver cirrhosis. As he pointed out, these results support the notion that train crews that are away from home regularly differ in their lifestyle from stationary workers which can spend their after-work hours at home with their families and friends. The increased lung cancer risk of mobile train crews might thus be mainly ascribable to lifestyle factors and in particular, smoking.

Lung cancer risk did not increase substantially in DE-exposed shop workers. However, as industrial hygiene measurements have shown, workers in locomotive shops generally had the highest levels of DE exposure among railroad workers (Woskie et al. 1988a, 1988b). Even if the selection of shop workers was, unfortunately, not specific to diesel locomotive shops (Laden et al. 2006), the fact that the lung cancer risk of at least partly exposed shop workers does not differ from that of the reference group conflicts with the hypothesized impact of DE on lung cancer risk. Based on these results, we can conclude that in addition to DE-exposure, a distinction between mobile and stationary railroad workplaces should be considered in the analysis of lung cancer risk in this cohort.

**Reanalysis of mortality rate ratios.**

The investigators reported that the lung cancer risk in DE-exposed jobs was increased in the birth cohort 1915–1919 [relative risk 1.49 (95% CI: 1.30–1.70)] and diminished consistently for older workers to merely 0.99 (95% CI: 0.75–1.30) in the birth cohort 1895–1899 (Table 2 in Garshick et al. 2004). Garshick and his colleagues interpreted this observation as evidence for a dose-response relationship between duration of DE-exposure and risk of lung cancer. In order to verify this relationship, but ignoring the adjustment for years of employment, we estimated mortality rate ratios (MRRs) using Poisson regression based on case numbers and person-years reported in Table 2 of the original paper (Garshick et al. 2004).

We specified the model

$$MRR(i,j) = \exp(\beta_0 + \beta_j + \gamma_i + \delta_{i,j}(2 - j)),$$

where $i$ denotes the job group, $j$ denotes the birth cohort (0 ÷ 1915–1919, …, 4 ÷ 1895–1899), $\beta$ is an age-specific parameter, $\gamma$ is the job specific parameter describing the age-adjusted relative risk with respect to the reference group, and $\delta$ is a log-linear term describing risk increase with increasing year of birth centered at the median birth cohort (per 5 years). The calculated MRRs are similar to that in the original paper, except that the original analysis resulted in a stronger divergence between youngest and oldest birth cohorts (Table 7). This difference, however, can be explained by the researchers’ adjustment for years of employment. Further, our calculations show that a significant risk increase over the birth cohorts is seen for engineers only: from 0.68 in the oldest birth cohort up to 1.45 in the youngest. In contrast, conductors have a steady MRR. It should also be noted that a corresponding test rejects the hypothesis of homogeneity between engineers and conductors with respect to trends in lung cancer MRR over the birth cohorts ($p = 0.02$).

The measurements of DE-exposure in railroad workers (Woskie et al. 1988a, 1988b) showed that DE-exposure for conductors was in general somewhat higher than for engineers. Therefore, it is surprising that a strong risk increase is observed in engineers but not in conductors.

Imputed smoking data for the whole cohort (Garshick et al. 2006) based on the case-control analysis (Garshick et al. 1987a) show that, according to information gathered from next-of-kin, engineers and conductors are fairly comparable with respect to tobacco consumption and, in addition, they smoked considerably more than the reference group. How does this fit with the very low MRR for engineers in the oldest birth cohort? The authors of the Finnish railway workers study suspected that changes in the work environment and the change in physical demands of work as an engineer have increased the possibilities for smoking during work and, hence, increased the risk for lung cancer (Nokso-Koivisto & Pukkala 1994). The results of the US railroad workers cohort seem to confirm this suspicion. Opportunities for engineers to smoke on a moving steam
Table 7. Impact of birth cohorts on mortality rate ratios in the Cohort study among US railroad workers.a

<table>
<thead>
<tr>
<th>Job group</th>
<th>Birth cohort</th>
<th>Estimates from Poisson model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1915–1919</td>
<td>MRR adjusted for age</td>
</tr>
<tr>
<td>Engineer</td>
<td>1.45</td>
<td>1.19 (1.07–1.31)</td>
</tr>
<tr>
<td>Conductor</td>
<td>1.35</td>
<td>1.28 (1.17–1.41)</td>
</tr>
<tr>
<td>Shop worker</td>
<td>1.12</td>
<td>1.07 (0.97–1.18)</td>
</tr>
<tr>
<td>Engineer &amp; Conductor</td>
<td>1.38</td>
<td>1.24 (1.14–1.35)</td>
</tr>
</tbody>
</table>

a Loglinear term describing risk increase with increasing year of birth, centered at the median birth cohort (calculated per 5 years).

In 2006 the first study results adjusted for smoking were published (Garshick et al. 2006). However, some critique on how the smoking data were assessed is warranted. As part of the nested case-control study, smoking history was assessed for workers of different birth cohorts that died in 1981 (i.e. assessed from their next-of-kin). This information was used to estimate the smoking habits of different birth cohorts for the cohort analysis (for birth cohort 1910 aged 49 in 1959, for birth cohort 1900 aged 59 in 1959, etc.). However, the lower life expectancy of smokers as compared to never-smokers resulted in an increasing underestimation of smoking habits with increasing age at death (Table 8). These data are therefore unsuitable for imputation into the cohort. Moreover, the original smoking information was gathered from next-of-kin only, which per se is prone to bias. For example, close relatives can estimate the smoking habits of their husbands or fathers only based on their memories about the situation at home. But, mobile workers such as train personnel could smoke at home fewer cigarettes than during periods away from home. Consequently, even a differential bias with respect to job group cannot be excluded.

Using the British doctors study as a reference (Doll et al. 2005), the low magnitude of risk estimates for different smoking categories in the case-control study (Garshick et al. 1987b) suggests that the smoking-related lung cancer risk in the railroad worker study was strongly underestimated.

**Summary.** In this large cohort study, lung cancer risk was correlated with certain job characteristics, in particular work location (mobile on the train versus stationary at a railway station). However, work location was correlated with lifestyle factors, such as smoking. In fact, difference in smoking habits between DE-exposed and non-exposed workers are a simple and probable explanation for observed increases in lung cancer risk. Secondly, the available information on smoking habits gathered by a case-control study is not comparable across birth-cohorts and, hence, the corresponding risk estimators for lung cancer, adjusted for smoking, also cannot be compared. Thirdly, the adjustment for years-worked is not justified and finally, our approximate re-calculations of the study data show that an increase of lung cancer risk is restricted to engineers only and this may have been caused by changes of smoking habits as a result of improved working conditions in the engine. Due to the strong likelihood for confounding bias of the lung cancer risk estimates, the study cannot add evidence for a causal link between DE-exposure and lung cancer risk and it cannot be used for QRA.
was a significant increase in the estimated lung cancer risk with length of employment as a long-haul trucker after 1959 (reported p-value for linear trend = 0.04). For workers with 18 and more years of service in this time period, an OR of 1.55 (95% CI: 0.97–2.47) was calculated.

To enable a quantitative exposure–response analysis, quantitative data on the exposure to elemental carbon (EC) were added to the database (Steenland et al. 1998). Data on exposure intensity are based on an industrial hygiene survey in the trucking industry from 1990 (Zaebst et al. 1991). Estimates for the time before 1990 were deduced using three different scenarios for decreasing diesel engine particulate emissions since 1970. The cumulative exposure to EC was further assumed to be proportional to the distance driven by heavy-duty trucks. Risk estimates were adjusted for age, race, smoking, diet, and asbestos. The investigators found a significantly increasing lung cancer risk with increasing cumulative exposure to EC. For the highest quartile of cumulative exposure, assuming the intermediate scenario and a 5-year exposure-lag, an OR of 1.64 (95% CI: 1.09–2.49) was calculated.

**Comments on the study**

The design of this study is very similar to that of the case-control study among deceased US railroad workers (Garshick et al. 1987a) and, hence, the limitations and methodological problems are similar. Unfortunately, the study reports contain only scarce descriptive data. No information is given about the distribution of age and smoking in the exposure-related subgroups.

**Validity of smoking information.** Information on smoking and exposure to asbestos was based exclusively on information from next-of-kin. Cigarette smoking information from next-of-kin is generally acceptable (McLaughlin et al. 1987). However, for drivers and in particular, long-haul drivers who spend most nights in their trucks, next-of-kin knowledge may not be accurate as their lifestyle on the road could be quite different from that at home. As the smoking prevalence among drivers is generally much higher than in other jobs (Jain et al. 2006; Sieber et al. 2014), residual confounding due to smoking is an issue in this analysis. Furthermore, the prevalence of ever-smoking decreases with increasing age at death in a fixed birth-cohort, due to the higher life expectancy of never-smokers. Hence, one should consider a bias in the smoking information, i.e. the smoking information is not comparable between birth cohorts. It can be assumed that in younger workers, who also have a longer exposure to DE, smoking has a stronger impact on the lung cancer risk than in older workers. Therefore, the residual confounding by smoking is stronger in younger workers. Moreover, substantial changes were observed with respect to smoking habits of the general population by birth cohorts, which are clearly visible in the mortality due to lung cancer data (Table 8). It cannot be assumed, that these changes were sufficiently reflected by the approximated smoking information gathered from next-of-kin. Hence, the adjustment for smoking is not adequate.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.21</td>
<td>5.39</td>
<td>7.41</td>
<td>9.59</td>
<td>11.73</td>
<td>11.81</td>
<td>11.06</td>
<td>8.82</td>
<td>5.95</td>
<td>3.72</td>
<td>2.09</td>
<td>1.32</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>3.21</td>
<td>5.39</td>
<td>7.41</td>
<td>9.59</td>
<td>11.73</td>
<td>11.81</td>
<td>11.06</td>
<td>8.82</td>
<td>5.95</td>
<td>3.72</td>
<td>2.09</td>
<td>1.32</td>
<td>0.81</td>
</tr>
</tbody>
</table>
| Data from CDC-WONDER (2009), Ribicoff and Terry (1961); own calculations. The reference values for study subjects involved in the US railroad worker study’s case-control analysis who were deceased in 1981 (Garshick et al. 1987a) are shaded.

when 1959 is used as the cut-point for the definition of exposure, all decedents who attained an age of at least 77 years belong to the lowest exposure tertile. Such a drastic displacement does not occur without the restriction of exposure to the time after 1959. The researchers emphasize that the lung cancer risk increased markedly with length of employment as a long-haul trucker after 1959 (OR = 1.55 for longest duration). However, for short-haul drivers an almost identical increase was observed. This is remarkable because a reverse trend is present for short-haul drivers without the use of the cut-point for the determination of the length of employment. Moreover, in contrast to the situation for long-haul trucks, the conversion to diesel-powered short-haul trucks did not start before the mid-70s and was completed between 1980 and 1992 (Garshick et al. 2012). Therefore, exposure of short-haul drivers to noticeable amounts of DE prior to death in 1982/1983 was unlikely. Consequently, it is more likely that lifestyle factors and smoking-related birth-cohort effects are responsible for the observed trends. This statement is supported by the findings for truck mechanics. No trend was observed with increasing years of employment in this job group with the highest DE exposure intensity in the whole study population.

**Summary.** This nested case-control study has similar methodological problems to the nested case-control study in US railroad workers discussed above. The close link between the quantification of DE exposure and birth-cohort, in conjunction with the substantial changes in the smoking habits by birth-cohort in the general population and the differential bias in the smoking data gathered from next-of-kin, leads to the conclusion that the findings of the study are not reliable. Consequently, the results are insignificant in view of the relationship between DE exposure and lung cancer risk and should not be used for QRA.

**Study in the unionized US trucking industry**

**Description of the study and authors’ main results**

The first analysis of this large-scale cohort study comprised more than 54,000 male unionized trucking industry
employees who had worked for at least one day in 1985 at one of the four participating companies (Laden et al. 2007). The mortality follow-up was performed from 1985 through 2000. Compared with the general US population, the investigators observed slightly elevated lung cancer mortality among male drivers [SMR = 1.10 (95% CI: 1.02–1.19)] and also among dockworkers. A considerably higher mortality was also observed for ischemic heart disease among drivers [SMR = 1.49 (95% CI: 1.40–1.59)], dockworkers and shop workers.

A more detailed analysis on lung cancer mortality was performed using a subset of this cohort, containing 31,135 male employees \( \geq 40 \) years of age in 1985 with at least 1 year of work in a trucking industry job (Garshick et al. 2008). Risk estimates for years of employment in exposed jobs were calculated with Cox models taking attained age as the time axis. In addition, the authors considered decade of age at study entry, calendar year, decade of hire, race, census region of residence and company in the regression model. In order to adjust for a HWSE, total years employed in the companies and years off work were included. Indirect adjustment for smoking was performed using information from an industrial survey in 2003 of active and retired trucking industry employees. HRs for 20 years of work in a job with DE exposure ranged from 1.40 (95% CI: 0.88–2.24) for long-haul truckers to 2.34 (95% CI: 1.42–3.83) for combination workers who performed dock and pickup and delivery (P&D) work.

A later analysis of the same data used a statistical model based on a national exposure assessment to estimate historical exposures to elemental carbon (EC) as a quantitative surrogate for the DE exposure in this cohort (Garshick et al. 2012). The estimation of historical REC exposures made use of REC measurements from 2001–2006 and accompanying job group-specific prediction models, developed to estimate REC based on a range of determinants. These prediction models were used to estimate historical exposure to REC, considering changes in work-related conditions over time based on a previous exposure assessment of the trucking industry (1988–1989) and changes in home terminal ambient REC based on a trend analysis of historical air pollution data (Zaebst et al. 1991; Davis et al. 2011).

The investigators yielded a HR for lung cancer of 1.07 (95% CI: 0.99–1.15) per 1 mg/m\(^3\)-months cumulative exposure to EC (excluding mechanics) with a 5-year exposure lag, adjusted for employment duration and other factors. In their discussion of the study’s limitations, the authors addressed the lack of individual information on smoking and the lack of pre-diesel occupational history, spanning on average over 10 years.

**Comments on the study**

**Adjustment for years worked.** The results of this large cohort study are primarily influenced by adjustments for total years employed by the companies and years off-work. The authors justify this adjustment as it controls for a supposed HWSE. However, a healthy-worker-survivor bias denotes a bias arising from the selection of unhealthy persons out of the workforce/cohort (Checkoway et al. 2004). The methodological issue of this study was rather a potential bias due to the selection of prevalent hires into the cohort known as left truncation (Applebaum et al. 2011).

The best way to avoid left truncation is to restrict the cohort to incident hires only. Yet, the trucker cohort consists of prevalent hires that have worked in the companies on average for 13 years prior to study entry. According to the published results, the duration of employment was correlated with the cumulative exposure to EC (0.55 < \( r < 0.74 \), depending on time lag) (Garshick et al. 2012). Hence, the parameter estimates for cumulative EC and years worked in the Cox model were mutually influential. Using the author’s preferred model (a 5-year exposure lag in the cohort excluding mechanics), the hazard ratio for lung cancer per year of work was 0.97 (95% CI: 0.96–0.99). In their second report, the investigators calculated the risk for various job groups for 20 years of work (Garshick et al. 2008). Estimates varied between 1.65 and 2.20 (without smoking adjustment). However, as 20 years of work corresponds to a hazard ratio of 0.54 (0.97\(^{0.05} \)), the overall lung cancer risk without adjustment for duration of work is actually not elevated. Indeed, HRs not adjusted for work duration reported in the latest publication were lower than the adjusted estimates and barely above the reference hazard (Garshick et al. 2012).

**Competing health-related risk factors.** It is clear that left truncation correlates to the health status of workers not included in the cohort, even though other reasons for leaving the job such as the opportunity to change to a better-paid job also exist. In general, left truncation induces bias only in the case when the cohort is heterogeneous in terms of the susceptibility to the effect of exposure. In this context, other health-related conditions and exposures should also be considered, especially in long-haul truck drivers that represent the largest occupational group among the DE exposed workforce.

According to the US National Health Interview Survey, the smoking prevalence among truck drivers is about 43% – one of the highest among 209 job groups in the years 1978, 1988 and 1990–94 (Lee et al. 2004). The 2010 National Survey of Long-Haul Truck Driver Health and Injury showed a prevalence of current smoking of 50.7% among long-haul truck drivers. This compares to the national average in the working population of 18.9% (Sieber et al. 2014). In addition, the US Long-Haul Driver Health and Injury survey demonstrated impressively that obesity and diabetes were twice as prevalent in long-haul truck drivers as in the 2010 US adult working population (obesity: 69% vs. 31%; diabetes: 14% vs. 7%, respectively) (Sieber et al. 2014; These et al. 2015). Further risk factors for drivers’ health including hypertension, high cholesterol, no physical activity, and fewer than 6 h sleep per 24-h period were identified. Therefore, it is not surprising that the mortality due to heart diseases is higher in long-haul truck drivers in comparison to other blue collar jobs, especially with respect to ischemic heart diseases (Robinson & Burnett 2005). Analogous results were seen in the unionized trucking industry study discussed here (Laden et al. 2007).

Further studies support the statement that truck drivers are at higher risk for lifestyle-related diseases (Hedberg et al.
1970 based on year of birth (models. Investigators assigned separate baseline hazards calculated from available records in the participating tion is linked to age at hire and thus with exposure estimates seems to be considerably underestimated. The underestima-
drivers, the largest subgroup in the cohort, the DE exposure must not be used in data analysis. Particularly for long-haul ment in the trucking industry. Consequently, this variable “total years on work” would therefore be falsely defined as 11 (instead of 44). Hence, the variable “total years on work” does not give reli-
early correlation between age at hire and total years worked. This was hypothesized to be due to pre-
dominant exposure to aged DE particles (Garshick et al. 2008). In our view the differences between mechanics and drivers with respect to other workplace-related factors, like the opportunity to smoke during working hours or lifestyle-related factors, seem to be a more plausible explanation for these findings. Hence, the exclusion of mechanics is not justified.

Conclusions on study results and recommendation for reanalysis. As a consequence of the methodological issues described above, only the results unadjusted for duration of work and including mechanics (Table 4 in Garshick et al. 2012) can be used to appraise the relationship between DE and lung cancer in this study in the trucking industry. Lung cancer risk was scarcely increased in ever-exposed workers, and the comparison of HRs by quartiles of various EC metrics does not show any trend. Consequently, the log-linear expo-
sure–response relationship (7% risk increase for each 1 mg/m³-month of cumulative EC), highlighted by the authors as the main result of their study, is clearly overestimated. A reanalysis of a sub-cohort, limiting the analysis to workers hired in the participating companies prior to their 25th birthday, could overcome the study limitations. To reduce the close correlation between years worked and cumulative DE exposure, one should include the years worked before diesel-
ization instead of the total years worked in the model.

Summary. The advantage of this cohort of workers in the trucking industry is the availability of quantitative exposure estimates. However, this advantage was leveled off due to the absence of information on exposures before employment in the participating companies. Particularly, long-haul drivers hired later in their working life may have already accumu-
lated several years of DE exposure in other companies. The adjustment variable “total years of work (in the participating companies)” was biased accordingly. Furthermore, as shown above, premature termination of the job in the trucking industry is related primarily to lifestyle-related factors (including smoking habits). Therefore, adjustment for “total years worked” was not indicated in this study. Due to the strong correlation between years worked and cumulative exposure, adjusting for “total years worked” in the preferred statistical model led to overadjustment bias; lung cancer risk in relation to DE was considerably overestimated. In conclusion, the study does not support a causal exposure–response relationship between DE and lung cancer and, due to the methodo-
logical shortcomings discussed, is not useable for QRA.
Study among bus garage workers in Stockholm

Description of the study and authors’ main results

The study comprises 695 male bus garage workers employed as mechanics, servicemen or hostlers for a minimum of six months in five bus garages in Stockholm between 1945 and 1970 (Gustavsson et al. 1990). Diesel-powered buses were first introduced in Stockholm in the beginning of the 1930s. The transition to diesel-powered buses was completed in 1945. Mortality (1952–1986) as well as cancer incidence, based on the data from the Swedish Cancer Registry (1958–1984), were investigated as study outcomes. Exposure intensity to DE and asbestos was estimated by a JEM, developed by industrial hygienists. Cumulative exposure was estimated by adding up the product of duration of exposure and exposure intensity over the time of employment.

The comparison of lung cancer mortality yielded a SMR of 1.22 (95% CI: 0.71–1.96; 17 cases) taking the occupational active male population of Stockholm as the reference. In a nested case-control study that included the 20 lung cancer cases (deceased and living cases combined) and six age-matched controls per case, an increase of lung cancer risk with increasing DE-index was estimated using a logistic regression model (OR = 1.37 (95% CI: 0.91–2.07) per 10 units of DE-index). For the highest exposure group, defined by a DE-index > 30, the relative risk estimated by weighted linear regression was 2.43 (95% CI: 1.32–4.47), taking the group with a DE-index < 10 as the reference.

Comments on the study

The cancer registry data showed 15 incident cases of lung cancer and 2 incident pleura mesothelioma cases. However, for the study period in question on average 30 cases of pleural mesothelioma were diagnosed in the Swedish male population annually with an increasing tendency (Järvholm et al. 1999). In contrast, the annual number of lung cancer cases, the number of pleura mesothelioma. From these data it can be assumed that the mortality due to mesothelioma is rather the number of pleura mesothelioma. Hence, the corresponding result should not be over-interpreted.

Summary. In conclusion, taking into account the lack of smoking data, non-adjustment for asbestos, probably weak asbestos data anyway, and the small number of lung cancer cases, the study can add only little evidence for the clarification of the relationship between DE and lung cancer risk.

Professional drivers cohort from London

Description of the study and authors’ main results

Based on the UK Health Service Central Register, the investigators identified 3392 men who were employed as professional drivers in London according to the 1939 census (Balarajan & McDowall 1988). The study was restricted to those drivers who needed special commercial licenses for their job, i.e. bus, coach, lorry, and taxi drivers. The mortality follow-up was conducted from 1950 to 1984.

The overall mortality was significantly reduced [SMR = 0.91 (95% CI: 0.87–0.95), own computations of 95% confidence intervals based on data reported in the publication], while the lung cancer mortality was clearly elevated [SMR = 1.47 (95% CI: 1.32–1.64)]. The highest lung cancer risk was observed among lorry drivers [SMR = 1.59 (95% CI: 1.41–1.79)], the lowest among taxi drivers [SMR = 0.86 (95% CI: 0.58–1.23)]. For lorry drivers, an elevated mortality risk due to stomach cancer [SMR = 1.41 (95% CI: 1.11–1.77)] and respiratory diseases [SMR = 1.17 (95% CI: 1.04–1.31)] was also reported.

Comments on the study

The major limitations of the study are the lack of information on smoking history and the assumption that the cohort members stayed in the same job until retirement. This assumption is weak, particularly when one considers the changes that occurred with respect to World War II. Even if one assumes that all cohort members stayed in their jobs over the study period, one would expect to see a greater lung cancer risk in taxi drivers compared to lorry drivers, as they drive mostly in the urban area of London and, hence, their exposure to DE should be higher. The fact that the risk difference between lorry and taxi drivers is also present for stomach cancer and nonmalignant respiratory diseases draws attention again to the differences between mobile and non-mobile workers with regard to differences in lifestyle habits.

Summary. The study lacks information on lifestyle factors such as smoking. However, the increase in the risk of lung cancer and other diseases seen in lorry drivers might be ascribable to such factors. Therefore, the study cannot contribute to answering the research question.

Professional drivers cohort from Geneva

Description of the study and authors’ main results

A retrospective study similar to the London cohort study of drivers was carried out in the Canton of Geneva. A total of
6630 male drivers with licenses for heavy-duty road goods vehicles, taxis, buses, coaches, and light-duty road passenger vehicles in 1949 or drivers getting new licenses between 1949 and 1961 were identified using the archives of the Cantonal licensing authorities (Gubaran et al. 1992). Mortality data from 1949 to 1986 was investigated. Cancer incidence from 1949 to 1988. Data on vital status and lung cancer mortality were gathered by linking records with the national population registry and the national death index. The cohort was compiled using membership rolls from the Truck Drivers’ Union and a cooperative taxi agency in Reykjavik. The follow-up of the cohort started in 1951 and ended in 1986. The cohort was divided into three subcohorts based on their noted occupation: professional drivers, nonprofessional drivers “more exposed” to exhaust gas due to their occupations, and “less exposed” nonprofessional drivers.

Lung cancer mortality was significantly elevated in professional drivers (SMR = 1.50 (90% CI: 1.23–1.81]) and somewhat lower in nonprofessional drivers (SMR = 1.32 and 1.21, respectively). In contrast to nonprofessional drivers, the professional group also showed a significantly elevated risk for stomach and rectum cancer, higher even than that for lung cancer. Furthermore, professional drivers had an increased risk of deaths due to esophageal cancer, circulatory disease and cirrhosis of liver.

Comments on the study
The cohort comprises a wide range of professions. Therefore, the comparison of occupational groups with respect to lung cancer without smoking information may be biased. Moreover, the mortality patterns of professional drivers show a strong impact of lifestyle-related factors.

Summary. This study lacks information necessary for the clarification of the relationship between DE and lung cancer risk.

Professional drivers cohort from Reykjavik, Iceland
Description of the study and authors’ main results
A retrospective cohort study was conducted in Reykjavik, Iceland, in order to compare truck and taxi drivers with respect to their lung cancer risk (Rafnsson & Gunnarsdottir 1991). The cohort was compiled using membership rolls from the Truck Drivers’ Union and a cooperative taxi agency in Reykjavik. The follow-up of the cohort started in 1951 and ended in 1988. Data on vital status and lung cancer mortality were gathered by linking records with the national population registry and the national death index. Almost all trucks have been powered by diesel engines since 1950.

For the 868 subjects who worked as truck drivers, a SMR of 2.14 (95% CI: 1.37–3.18) was calculated. This was based on 24 cases of cancer of the lung, bronchus or trachea classified by ICD-7, and, thus, also included pleura mesothelioma. Among taxi drivers, there were 12 cases with a resulting lower SMR [SMR = 1.39 (95% CI: 0.72–2.43)]. The lung cancer risk was not affected by the duration of employment.

Comments on the study
According to a survey cited in the paper, truckers considered themselves significantly less likely to be never-smokers compared to both taxi drivers and all participants of the survey combined (11.9% vs. 22.5% or 22.8%, respectively). This difference might partly explain the study results. Yet, mortality due to nonmalignant respiratory diseases was unusually low among truckers [SMR = 0.50 (95% CI: 0.28–0.82), based on 15 cases], in contrast to taxi drivers. The SMR for malignant and nonmalignant respiratory diseases combined was also less than one.

In addition, the authors reported the SMR by the duration of follow-up up to ≥50 years (Table 7 in Rafnsson & Gunnarsdottir 1991). However, duration of follow-up for each subject, by definition, could not exceed 38 years. It therefore seems probable that the authors designated the time since hire (TSH) as follow-up duration. Truckers with less than 20 years since hire had the highest risk observed [SMR = 4.35 (95% CI: 1.18–11.13), based on four incident cases (own calculations)]. For the categories 20–39 years and >40 years since hire, the lung cancer risk was somewhat lower [SMR = 2.64 (95% CI: 1.54–4.23), 17 cases, and SMR = 2.31 (95% CI: 0.48–6.74), 3 cases, respectively]. Unfortunately, the risk distribution by calendar time was not reported.

In their paper the investigators refer to findings from a Swedish study (Ahlberg et al. 1981) saying “It has been suggested that smoking among drivers could be regarded as a special case because of the lack of air circulation in their vehicles, the result being that drivers are subject to more indirect exposure from tobacco smoke than are others who smoke on the job”. Therefore, smoking could be a potential reason for the elevated lung cancer mortality. However, systematic bias in the documentation of the underlying cause of death cannot be excluded. Further complicating the findings is that truck drivers in Iceland usually repaired and maintained their own trucks and, thus, were exposed to other known risk factors for lung cancer and mesothelioma, such as asbestos.

Taxi passengers were allowed to smoke in taxi cabs in Iceland until ~1980. This fact may help explain the somewhat elevated lung cancer risk in taxi drivers while the smoking habits in this job group were almost identical to that of the survey’s reference group. Furthermore, taxi drivers did not usually repair and maintain their own vehicles, reducing their exposure to asbestos and other potentially hazardous substances.

Summary. Due to likely confounding of the lung cancer risk estimates by asbestos exposure and smoking, the study does not improve our understanding of the DE-lung cancer hypothesis.

DE-exposed workers in Finland
Description of the study and authors’ main results
In a registry-based study, Finnish census data from 1970 was linked with the Finnish Cancer Registry from 1971 to 1995 (Guo et al. 2004). Exposure to DE and gasoline engine exhaust was assessed according to the occupation held for the longest period in the census, in combination with the FINJEM job exposure matrix (Kauppinen et al. 1998). The FINJEM is based
on available exposure measurements and the judgment of 20 Finnish occupational hygienists. It covers major occupational exposures in Finland since 1945 by occupation and calendar time. Cumulative exposure to DE was assessed according to the concentration of nitrogen dioxide (NO₂) in the air (mg/m³) multiplied by time in years. In addition, FINJEM contains a rough estimate of smoking prevalence (proportion of daily smokers). Risk estimates in the study were presented as standardized incidence ratios (SIR) by job codes, or as risk ratios derived from internal comparisons by Poisson regression, adjusted for smoking, asbestos, quartz dust, socio-economic status, age, and period of follow-up.

The lung cancer risk for male bus, taxi, and truck drivers was SIR = 0.89 (95% CI: 0.78–1.00), SIR = 1.10 (95% CI: 0.96–1.26), and SIR = 1.13 (95% CI: 1.04–1.22), respectively. Internal comparison did not show an increase in lung cancer risk for males by increasing cumulative exposure to DE. For the highest exposure category (>10 mg/m³-years) a RR of 0.95 (95% CI: 0.83–1.10) was calculated. The risk pattern for males with respect to gasoline exhaust was similar. In contrast to these results, a significant risk increase for lung cancer with increasing exposure to gasoline engine exhaust was observed for females, but not with respect to DE.

**Comments on the study**
Strengths of the study include the large database, the opportunity to link the data with the cancer registry and the availability of smoking data albeit as rough approximates. The investigators wrote “without adjustment for asbestos or quartz dust, or smoking, our results would also have shown an effect of DE on lung cancer risk”. This implies that the adjustment for smoking had a noticeable impact on the risk estimates.

For the quantification of DE-exposure the FINJEM was used (Kauppinen et al. 1998). This JEM used the concentration of nitrogen dioxide (NO₂) in air (mg/m³) as the indicator for DE exposure. Job-specific exposure values were weighted by the proportion of exposed subjects in a job group and summed up over the tenure to estimate cumulative exposure. Further assumptions were that exposure started at age 20 and that workers continued to work in the same job as recorded in the 1970 Census. Hence, the deviation of the exposure estimate from the actual exposure for a worker might be greater in this study than in other cohort studies based on quantitative exposure estimates, i.e. a more severe exposure misclassification must be considered. This effect often leads to shrinkage of the risk estimates toward one.

**Summary.** The study did not show a relationship between DE-exposure and lung cancer and thus does not support the hypothesis of a causal link. In view of the probably high level of quantitative exposure misclassification the study should not be used for QRA.

**Cohort of bus drivers and tramway employees in Copenhagen**

**Description of the study and authors’ main results**
A retrospective cohort study was conducted among 18,174 bus drivers and tramway employees who worked between 1900 and 1994 for a large public transportation company in Copenhagen (Soll-Johanning et al. 1998). Employment records were obtained from the company files and linked with cancer incidence data from the Danish Cancer Registry, 1943–1992. Information on smoking was not available. Using national reference rates, lung cancer risk was found to be significantly elevated in male and female employees (SIR = 1.6 (95% CI: 1.5–1.8), 473 cases, and SIR = 2.6 (95% CI: 1.5–4.3), 15 cases, respectively). A trend by duration of employment could not be observed.

To overcome the shortcomings of the cohort study, in addition the investigators carried out a nested case-control study. Cases were restricted to those with a personal identification number and, thus, employees still alive in 1968 when the Central Population Registry was established (Soll-Johanning et al. 2003). Cases older than 85 years and deceased cases without a contactable next-of-kin were also excluded. In total 153 lung cancer cases and 351 randomly sampled controls, matched by year of birth and vital status, were included in the analysis. All controls were lung cancer free at the time of the incident cancer diagnosis of their matched cases. Potential controls that died from cancer or nonmalignant respiratory diseases were also excluded. Information about smoking habits and occupational history were obtained from cases or next-of-kin. 98% of cases and 94% of controls were classified as smokers; almost all smokers had a smoking history of more than 20 years. The investigators reported that, when adjusted for smoking, being a bus driver for more than 20 years was associated with a decrease in lung cancer risk. Additionally, a decreasing trend in the smoking adjusted odds ratios was observed [OR = 0.97 (95% CI: 0.96–0.99) per year employed]. Ever employment in other driving jobs or as a mechanic with DE exposure (37% and 35% for cases and controls, respectively) was also inversely associated with lung cancer risk.

**Comments on the study**
The primary study aim was to analyze the impact of air pollution on lung cancer risk. Consequently, a specific analysis for DE-exposed jobs was not performed. However, assuming that DE exposure is higher for drivers of DE-powered vehicles, as was done in other studies, a further differentiation between bus drivers and tramway employees would have been helpful, as tramways are powered electrically.

The slightly increased lung cancer risk for transportation workers when compared with the general population could be explained by smoking. As the case-control analysis showed, almost all included subjects were smokers.

**Summary.** The study results are non-conclusive with respect to a relationship between DE and lung cancer risk.

**Cohort study among urban bus drivers in three Danish cities**

**Description of the study and authors’ main results**
The cohort study included 2037 male urban bus drivers in the three largest Danish cities with follow-up for cancer
incidence from 1979 to 2003 (Petersen et al. 2010). An earlier report covered cancer incidence until 1984 (Netterstrom 1988). Mailed questionnaires were used to collect information on employment as a driver prior to the start of the follow-up period. Employment as a professional driver thereafter was ascertained using the Supplementary Pension Fund. Membership of the fund is compulsory for all wage earners in Denmark. The Danish Cancer Registry was used to identify cancer cases. The combined cancer incidence rates for the general male population of the three cities were used for external comparison. The SIR for cancer incidence until 2003 was 1.09 (95% CI 1.0–1.2), whereas risks for lung cancer [SIR: 1.2 (95% CI 1.0–1.4)] and bladder cancer [SIR: 1.6 (95% CI 1.2–2.0)] were somewhat increased. 70% of the participating drivers reported active smoking at the start of follow-up. The internal Cox regression analysis adjusted for smoking status (never, former, current), amongst others, did not show an association between years of work as a professional driver and lung cancer. This was true for both the questionnaire-derived employment information and the complete employment information derived from the pension fund registry. For bladder cancer, the longest duration (>25 years) of work as a professional driver was associated with a non-significantly increased hazard ratio of 1.31 (95% CI: 0.70–2.48).

Comments on the study
The study authors concluded that the increased SIRs from the external analysis were probably due to the higher smoking prevalence among professional drivers, compared with the reference population. Self-reported smoking prevalence amongst the bus drivers was higher than the prevalence reported in a separate study of the general male population of Copenhagen, conducted at a similar time (70% versus 60%, respectively). The authors further discuss that any true small risk increases might not have been detected as a result of small study power or because employment years as a driver might not act as a good surrogate for cumulative exposure to traffic pollutants. On the other hand, as smoking was only considered in terms of three crude categories and derived from participants’ smoking history prior to the study follow-up period, the observed study results are probably still biased due to residual confounding by smoking.

Summary. The study on public bus drivers in three Danish cities does not support the hypothesis of a causal association between DE exposure and lung cancer. Residual confounding of risk estimates due to smoking is likely. Not least because of missing quantitative data on exposure the study cannot be used for QRA.

Study on Danish truck drivers

Description of the study and authors’ main results
This cohort study comprised 14,225 truck drivers and 43,024 other unskilled male laborers, aged between 15 and 74 years and occupationally active on the day of the Danish census in 1970 (Hansen 1993). The cohort was followed for exactly 10 years by a record linkage between the 1970 census register, the Central Population Register 1970–1980, and the Danish Death Certificate Register 1970–1980. A total of 627 drivers and 3811 non-drivers died during the study period. Individual information on smoking habits was not available. Non-drivers were used as the reference population assuming that the smoking habits were comparable between groups.

The mortality from lung cancer was significantly elevated in the drivers [SMR = 1.60 (95% CI: 1.26–2.00), 76 cases]. However, elevated risks were seen also for laryngeal cancer and other respiratory cancers [SMR = 1.40 (95% CI: 0.29–4.10), 3 cases; SMR = 1.70 (95% CI: 0.55–3.97), 5 cases; respectively]. The study authors noted the missing information on smoking, the short follow-up period, and the exposure definition based only on the active job at the reference date as the main limitations of their study design.

Comments on the study
The age structure between the two cohorts was very different. Only 7% of truck drivers were over 55 years of age, in comparison to 24% of the reference group. One explanation for this difference could be the rapid rise of the trucking industry in Denmark following WWII. Assuming this to be true, the authors’ assumption that “most Danish truck drivers remain truck drivers throughout their working lives” is not unreasonable, at least for the younger age strata. However, for the truckers older than 55 years in 1970 (>30 years in 1945), this assumption is incorrect. This cohort of men may have held other jobs before 1945 most probably as unskilled laborers – with unknown occupational exposures. The mortality data reflect the risk difference between the two age strata. For truckers up to 55 years of age the SMR is 1.07 (95% CI: 0.71–1.53, 29 cases) but for older truckers our calculation, based on data from Table 5 in Hansen (1993), yields an SMR of 2.32 (95% CI: 1.70–3.08, 47 cases).

Summary. The conclusiveness of the study is very limited, taking into account the shortcomings described.

Cohort study among Swedish dock workers

Description of the study and authors’ main results
In a retrospective cohort study, the mortality and cancer incidence of 6071 Swedish dock workers in the period from 1961 to 1980 was investigated (Gustafsson et al. 1986). The cohort comprises all dock workers that were employed between 1961 and 1973 at Swedish ports for at least six months. Diesel-powered trucks were introduced to Swedish ports in the late 1950s, and their number increased rapidly in the early 1960s. Data on mortality and cancer incidence were obtained by record linkage of the cohort data with the Cause-of-Death Registry as well as the Cancer Registry. The overall mortality was significantly lower than that for the reference population [SMR = 0.89 (95% CI: 0.84–0.94)], but the mortality due to lung cancer was significantly elevated [SMR = 1.32 (95% CI: 1.05–1.66); 70 cases]. Based on incidence...
data, the investigators calculated an even higher lung cancer risk \( \text{SIR} = 1.68 \) (95% CI: 1.36–2.07); 86 cases).

Later, a nested case-control study was conducted (Emmelin et al. 1993). It was planned to include the 20 ports at which lung cancer cases had occurred. Due to the lack of exposure data in some ports, the study was further restricted to 15 ports where dieselization of dock work started in 1957 and was finalized in 1963. The baseline cohort included 6573 men employed as dock workers for at least 6 months between 1950 and 1974. Up to four controls were matched to each eligible lung cancer case diagnosed between 1960 and 1982 by date of birth and port. In total, 50 cases and 154 controls were included in the study. Information on employment was collected from company and union records. Smoking data were based on information from living controls, next-of-kin of the deceased and colleagues of the study subjects. The three following variables have been used as surrogate markers for exposure to DE: (a) time worked with diesel equipment ("machine time") (b) cumulative fuel consumption (calculated as liter diesel per capita), and (c) a cumulative "exposed time" variable where "exposed time" was equal to 1 for years in which the annual fuel per capita exceeded the lower quartile of all years/all ports distribution, and a value of 0 was given for the other years. The three DE exposure variables were separately included in the analysis, each with three categories (low, medium, high).

Among the three exposure variables, the highest risk estimate was calculated based on the exposed time. The odds ratio for high versus low exposure was 2.9 (90% CI: 0.8–10.7). Exposed time was then used for further analysis. It was shown that the lung cancer OR of high DE-exposed smokers increased to 28.9 (90% CI: 3.5–240), using low DE-exposed nonsmokers as the reference group. The investigators concluded from their results that DE has an independent effect on the lung cancer risk and that a strong interaction exists between the effects of smoking and DE.

**Comments on the study**

**Cohort analysis.** The basic analysis (Table 3 in Gustafsson et al. 1986) shows a considerable heterogeneity between port regions. Applying the tools for meta-analysis it appears that the measure \( I^2 \) describing the variation in the SMRs attributable to heterogeneity is 61% (\( p = 0.002 \)). Hence, the lung cancer risk is heterogeneous between port regions. The highest risk is seen in the county Gävleborg on the Baltic coast in the middle of Sweden. Here, the lung cancer risk was 2.76 times higher than for dock workers in Göteborg, the port with the largest workforce included in the study (\( p < 0.005 \); 95% CI: 1.32–5.55; own calculations). It is worth mentioning that Gävleborg is located in the vicinity of the mining region Bergslagen, where copper, iron, and lead ore have been mined. Therefore, it is likely that the type of goods shipped from the port is also an important cofactor in the analysis.

Smoking prevalence among Swedish men is considerably lower than in other western-European countries (Pierce 1989). In addition, smoking prevalence is the highest in the birth cohorts from 1942 to 1949 and considerably lower in men born earlier (Midlöv et al. 2014). Therefore, we would assume that for this study, a smoking prevalence in the reference population of \( \sim 35\% \) is realistic. The case-control study nested into this cohort recorded a smoking prevalence of 61% among control subjects. The marked difference between the smoking prevalence in the reference population and that in dock workers could on its own explain the higher lung cancer risk observed in the dock workers cohort.

**Case-control analysis.** The investigators focused their case-control study primarily on the exposed time, an intensity-weighted sum of exposed months, as the exposure variable. As seen in Figure 1 of the paper by Emmelin et al. (1993), the annual fuel consumption increases with calendar time, which is plausible assuming an increase in productivity. Hence, more recent years are more likely to be classified as exposed years and younger dock workers (with probably riskier smoking habits) would accrue more exposed years than older workers. Unfortunately, the authors also did not verify whether the surrogate measures for DE exposure in a port are correlated with the lung cancer SMR in that port.

The inclusion of smoking information in the model as a binary variable leads to considerable residual confounding, particularly in view of the described procedure of categorizing former smokers as either "nonsmokers" or "smokers" based on whether they had stopped smoking 5 years prior to the incident cancer case’s date of diagnosis.

Using the published frequencies of cases and controls, we re-analyzed the data to verify the study’s results for exposed time in combination with smoking (Table 5 in Emmelin et al. 1993). We applied unconditional logistic regression as we did not have access to the original database necessary for the conditional approach. In contrast to the authors’ results, our analysis yielded much lower risk estimates. The original publication found an OR of 2.9 for high DE-exposed nonsmokers and an OR of 28.9 for high DE-exposed smokers. Our corresponding results were 1.11 and 6.87, respectively. The comparison of various models by AIC and BIC shows that smoking is the only important variable in the dataset (Table 9). The OR for the binary smoking variable was 4.68 (95% CI: 1.88–11.66), i.e. close to the results of the investigators’ conditional logistic model for this variable. A likelihood-ratio test (i.e. without penalizing additional parameters) indicates that the interaction term between DE-exposure and smoking is not statistically significant (\( p > 0.8 \)). The corresponding ORs for the medium- and high-DE exposure categories were 1.35 (95% CI: 0.58–3.12) and 1.83 (95% CI: 0.78–4.29), respectively. A further likelihood-ratio test even rejected the inclusion of DE-exposure variable into the model. The large discrepancy in results seen between the original conditional and unconditional logistic regression models highlight that the results of our unconditional analysis could have considerably underestimated the actual lung cancer risk. However, the much larger risk estimates for the conditional approach suggest that the case-control matching variables "age" and "port" play an important role in the interpretation of the original results and, thus, should be considered together with smoking differences as sources of confounding bias.
Values for the best fitting model are given in bold. Exposure data, it is not suitable for QRA. Between exposure to DE and the risk of lung cancer. Due to missing information on DE exposure intensity and detailed information on smoking. Nevertheless, in view of the non-elevated risk in the combined group of DE-exposed workers (SIR = 1.04 (95% CI: 0.86–1.24), own computation based on the information in the study report), the study does not indicate a causal link between DE exposure and lung cancer. Due to missing quantitative data on exposure, the study data are not useable for QRA.

Cohort Study in the US construction industry

Description of the study and authors’ main results

This retrospective, mortality cohort study consisted of 34,156 male Heavy Construction Equipment Operators Union members with at least 12 months of membership between 1964 and 1978 (Wong et al. 1985). Overall, 9.8% of the cohort died before 31 December 1978, the end of the follow-up period. For a further 5.2% the vital status could not be determined.

The analysis of the entire cohort yielded an unremarkable risk estimate for lung cancer mortality [SMR = 0.99 (95% CI: 0.88–1.10)]. An internal comparison revealed an increasing risk of lung cancer with longer duration of union membership. Workers with at least 20 years of membership had a SMR of 1.07 (95% CI: 0.91–1.25; own computation based on numbers given in study report), while workers with less than 5 years of membership had a SMR of 0.45 (95% CI: 0.22–0.83). A similar result was obtained in the internal analysis by latency. Half of the lung cancer cases were observed among already retired union members, and the risk estimate was considerably higher in this stratum (SMR = 1.64 (95% CI: 1.39–1.92)).
Comments on the study
The lung cancer mortality for the cohort was virtually identical to the reference population. An excess lung cancer risk was seen only amongst retired union members. Subjects in this group were coming to the end of their working lives at the time of the study start. The authors noted "it has been reported that members of the union historically have been exposed to concentrations of respirable dust substantially in excess of permissible levels". The significantly elevated mortality from emphysema among union workers as compared to the general population supports this suspicion [SMR = 1.65 (95% CI 1.36–1.98)]. Moreover, emphysema mortality considerably increased with the duration of union membership. When the original cohort was restricted to retired union members, in addition to the elevated lung cancer mortality an increased SMR is also seen for emphysema (SMR = 2.77), liver cancer (SMR = 2.07), liver cirrhosis (SMR = 1.74), and cancer of the digestive system (SMR = 1.42). This implies that lifestyle factors, including smoking, alcohol consumption, and dietary habits, play an important role in the increased disease risk seen in the older cohort. Hence, it is questionable if DE exposure, expressed by duration of union membership, could really be responsible for the excess lung cancer risk observed.

Individualized data on smoking was not available for the cohort. As the authors reported, a small-scale survey indicated that 27 (25.2%) of the 107 participating union members were never-smokers, which is only slightly lower than the reference value from the 1970 National Center for Health Statistics’ Interview Survey (30.7%). However, these data are not sufficient to compare, for example, the smoking habits by duration of membership in the union. Furthermore, the survey was conducted in union members that were still active. It may, therefore, incorrectly describe smoking habits of inactive, retired or deceased workers. The study report includes the information that at the end of follow-up in 1978 only 46.1% of the cohort members were still active union members. Other relevant information about this survey, especially the response rate, was also not provided.

Summary. The main limitations of the study are that only partial work histories were available for the cohort and the lack of data on DE-exposure levels. Hence, even semi-quantitative exposure estimates were not available. The mortality pattern for retired union members suggests that primarily lifestyle factors such as smoking and alcohol consumption are responsible for elevated risk estimates. Other occupational exposures such as silica dust may have contributed to the increased lung cancer risk. The study is not-informative with respect to a possible causal relationship between DE-exposure and lung cancer risk and also does not allow QRA.

Excluded cohort studies
Several studies were excluded from detailed discussion due to insufficient data on the estimation of an exposure–response-relationship between DE and lung cancer risk. The reasons for exclusion are briefly described below.

Occupational differences in risks of lung cancer were described based on data from the Los Angeles County Cancer Surveillance Program (Menck & Henderson 1976). Despite the cancer data being of high quality, the lack of specific data on DE exposure, as well as on duration of exposure and smoking habits, do not permit any statement about the impact of DE on lung cancer risk.

The large-scale Cancer Prevention Study II, organized by the American Cancer Society, was used as the study base for an investigation of DE effects on mortality (Boffetta et al. 1998). However, this analysis includes information on the first two years of follow-up only (1982–1984). Meanwhile the mortality follow-up for this large cohort (1.2 million men and woman) has been extended to 2006. Unfortunately, we were unable to find a reanalysis of the current data on DE and mortality. Consequently, the results can be regarded as preliminary only.

The Minnesota Highway Maintenance Worker Study (Bender et al. 1989) lacks detailed information on lung cancer mortality. The lung cancer SMR was 0.69 (95% CI: 0.52–0.90). Further details were reported only for cancer sites with elevated SMRs.

Members of a large prepaid health plan formed the basis for another cohort study (Van Den Eeden & Friedman 1993). However, the analysis was based only on self-reported occupational exposure to engine exhaust. A specific analysis for DE was not possible, as was any differentiation by duration of exposure.

Based on the Swedish Cancer Environment Register III, the cancer incidence during 1971–1989 was investigated (Boffetta et al. 2001). Probability and intensity of DE exposure in 1960 and 1970 was assigned by a job-exposure-matrix. Information on duration of exposure was missing, as well as data on smoking.

A further large-scale cohort study involved subjects that were at any time between 1989 and 2004 members of a trade association that provides services to truck drivers (Birdsey et al. 2010). Data on employment history, exposure to DE and individual smoking habits are not available.

Hospital-based case-control studies
Overview of the studies
The hospital-based case-control studies are summarized in Tables 10 and 11. They include a study in 18 hospitals across six US cities (Boffetta et al. 1990) and a study from Montreal, Canada, with a hospital- and population-based control group (Parent et al. 2007) (Table 10). Also included are several component studies of a European, multi-center, case-control project financially supported by the INCO-Copernicus program of the European Union (Olsson et al. 2011), and three studies conducted in France and Italy (“LUCA”, “PARIS”, “ROME”) which, together with the INCO-studies, contributed to a pooled analysis of European and Canadian case-control studies (Olsson et al. 2011) (Table 11). The lung cancer ORs for the highest category of DE-exposure ranged between 0.95 (95% CI: 0.52–1.74) in the French LUCA study and 1.76 (95% CI: 0.80–3.90) in the PARIS study, also conducted in France (Tables 10 and 11).
Table 10. Hospital-based case-control studies.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study region and subjects</th>
<th>Recruitment of cases</th>
<th>Sampling of controls</th>
<th>Proxy respondents in cases/controls</th>
<th>Adjustment for smoking(^a)</th>
<th>Adjustment for SES</th>
<th>Definition of exposure</th>
<th>Full adjusted OR (95% CI) for highest exposure category</th>
<th>Trend with increasing exposure(^c)</th>
<th>Considered in IARC's 2012 decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boffetta et al. (1990)</td>
<td>6 US cities</td>
<td>1977–1987; 1985–1987</td>
<td>Exclusion of tobacco-related diseases</td>
<td>0%</td>
<td>3(^b)</td>
<td></td>
<td>Years of education</td>
<td>1.49 (0.72–3.11)</td>
<td>(p = 0.18) Yes</td>
<td></td>
</tr>
<tr>
<td>Parent et al. (2007)</td>
<td>Canada, Montreal area; men</td>
<td>1979–85</td>
<td>Cancer despite lung cancer</td>
<td>29.4%/19.2%</td>
<td>4(^c)</td>
<td>Family income</td>
<td>Duration of DE exposure</td>
<td>1.17 (0.40–3.41)</td>
<td>(p = 0.25) No</td>
<td></td>
</tr>
</tbody>
</table>

SES: socio-economic status.

\(^{a}\)Smoking habits: 0 – no adjustment, 1 – adjustment for status (never, former, current), 2 – adjustment for status cross-classified by further information, 3 – packyears, 4 – packyears plus time since quitting.

\(^{b}\)Adjustment for average quantity of cigarettes/day (continuous); this seems to be not straightforward as the actual 8-level-definition of smoking in this study involves current smokers (1–20, 21–40, 41+ cigarettes/day), ex-smokers (1–20, 20–40, 40+ cigarettes per day), pipe and/or cigar smokers, and non-smokers.

\(^{c}\)No: trend not detected; \(p\)-value given if reported.

Overview of the studies

The population-based case-control studies on DE and lung cancer are summarized in Table 12. Two studies were conducted in Canada. The first study, from Montreal, ascertained cases via incident hospital diagnoses. A population-based control group was included in Canada. The first study from Toronto was conducted in Canada. The first study, from Montreal, ascertained cases via incident hospital diagnoses. A population-based control group was included (Parent et al. 2007). The initial analysis using incident cases from 1979–1985 (Parent et al. 2007) was followed by another analysis using the entire period of study (Olsson et al. 2011). The initial analysis used a control group with the (hypothetical) source population in regards to the association between exposure and outcome of interest. In this case, excluding tobacco-related diseases ensures better comparability of the hospital control cases with the (hypothetical) source population in regards to the association between exposure and outcome of interest. In this case, excluding tobacco-related diseases from the set of permissible diseases among controls. This would be an appropriate approach for studies with smoking as the exposure of interest. The other hospital-based studies other than lung cancer (Parent et al. 2007). The other hospital-based studies other than lung cancer (Parent et al. 2007). The other hospital-based studies other than lung cancer (Parent et al. 2007). The other hospital-based studies other than lung cancer (Parent et al. 2007). The other hospital-based studies other than lung cancer (Parent et al. 2007). The other hospital-based studies other than lung cancer (Parent et al. 2007).
Table 11. Pooled analysis of European and Canadian case-control studies (Olsson et al. 2011; Straif et al. 2010; further indicated references reported about type of controls).

<table>
<thead>
<tr>
<th>Part of analysis</th>
<th>Type of study/controls</th>
<th>OR estimates (95% CI) for highest cumulative exposure category</th>
<th>Considered in IARC’s 2012 decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete pooled analysis</td>
<td>Olsson et al. (2011)</td>
<td>1.31 (1.19–1.43)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Straif et al. (2010)</td>
<td>1.27 (1.14–1.41)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.14 (1.03–1.26), adjusted for education</td>
<td></td>
</tr>
<tr>
<td>Component studies of pooled analysis Olsson et al. (2011)</td>
<td>INCO studies (Zeka et al. 2006; Cassidy et al. 2007)</td>
<td>Czech Republic H (excl. cancer and TRD) 1.16 (0.64–2.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hungary H (excl. cancer and TRD) 1.27 (0.76–2.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poland H (excl. cancer and TRD), P (PR) 1.77 (1.08–2.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Romania H (excl. cancer and TRD) 0.99 (0.40–2.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Russia H (excl. cancer and TRD) 1.17 (0.74–1.86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Slovakia H (excl. cancer and TRD) 1.60 (0.80–3.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>UK P (HC) 0.93 (0.59–1.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separate studies on DE and lung cancer (see also Table 12)</td>
<td>AUT (Germany), (Brüske-Hohlfeld et al. 1999)</td>
<td>P (PR, RDD) 1.67 (1.38–2.03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HDA (Germany), (Brüske-Hohlfeld et al. 1999)</td>
<td>P (PR) 1.20 (0.87–1.66)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EAGLE (Italy), (De Matteis et al. 2012)</td>
<td>P (HC) 0.98 (0.75–1.27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TURIN/VENETO (Italy), (Richiardi et al. 2004; Richiardi et al. 2006)</td>
<td>P (PR) 1.18 (0.89–1.57)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LUCAS (Sweden), (Gustavsson et al. 2000)</td>
<td>P (PR; two control groups, of which one is matched to cases regarding vital status in 1990 (with exclusion of smoking-related causes of death)) 1.20 (0.91–1.59)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MONTREAL (Canada), (Pintos et al. 2012)</td>
<td>P (ER) 1.36 (0.96–1.91)</td>
<td></td>
</tr>
<tr>
<td>Additional studies</td>
<td>MORGEN (The Netherlands)</td>
<td>Nested in a prospective cohort study 0.62 (0.05–7.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LUCA (France), (Stücker et al. 1995; Stücker et al. 1999; Stücker et al. 2002)</td>
<td>H (excl. cancer and absence of non-cancerous lung diseases among controls) 0.95 (0.52–1.74)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PARIS (France), (Boffetta et al. 1998)</td>
<td>H (excl. TRD) 1.76 (0.80–3.90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ROME (Italy), (Bochicchio et al. 2005)</td>
<td>H (excl. TRD and diseases related to dietary habits) 1.23 (0.74–2.06)</td>
<td></td>
</tr>
</tbody>
</table>


In addition to the Canadian studies, a study was conducted in New Mexico, including cases from 1980–82 compiled from the New Mexico cancer registry (Lerchen et al. 1987). This was the oldest study included in our review. We also reviewed two studies conducted in the city of Turin (Richiardi et al. 2006) and the Lombardy region of Italy (De Matteis et al. 2012), three in France (Menvielle et al. 2003; Wild et al. 2012; Matrat et al. 2015), one in Germany (Brüske-Hohlfeld et al. 1999), one in Stockholm, Sweden (Gustavsson et al. 2000), and one study with two reports (one of which was restricted to nonsmokers) in Hong Kong, China (Tse et al. 2009, 2011). Furthermore, a pooled analysis of case-control studies in Europe and Canada (Olsson et al. 2011) included data from several of the aforementioned studies (Table 11): the German study with two study centers (Brüske-Hohlfeld et al. 1999), the Italian study in the Lombardy region (EAGLE study) (De Matteis et al. 2012), the study from Turin, Italy, which extended to a second study center in the Venetian region (Richiardi et al. 2004, 2006), and the Swedish study in Stockholm (Gustavsson et al. 2000). Additional population-based case-control data involved in the pooled analysis came from the Liverpool and Warsaw centers of the INCO Copernicus project (Cassidy et al. 2007) and, as a special case, from the Netherlands sub-cohort of the European Prospective Investigation into Cancer and Nutrition (Table 11).

The lung cancer odds ratios for subjects with the highest cumulative DE-exposure or longest exposure duration ranged from 0.7 (Menvielle et al. 2003) to 1.8 (Pintos et al. 2012). Control subjects were recruited using population registries, electoral rolls, health-related population databases, random-digit-dialing procedures or, in one case, property assessments data. Adjustment for smoking varied and included information on smoking status plus duration or intensity of use, pack-years or pack-years plus further information such as time since quitting. In further sections, we aim to describe special methodological features of population-based case-control studies on DE-exposure and lung cancer, which are important for the interpretation of the relationships under question.

**Population-based case-control study in Stockholm, Sweden**

**Description of the study and authors’ main results.** In a study conducted in Stockholm, Sweden, which ascertained
lung cancer cases from the national cancer registry between 1985 and 1990 (Gustavsson et al. 2000), two control groups were selected from the Stockholm population registry from the same years. The first control group was selected among subjects that were alive at the end of the calendar year in which a case, matched for age, arose. The second "mortality-matched" control group was selected among subjects alive at the start of the calendar year in which a case arose and in addition, case- and control-subjects were matched with regard to vital status as of 31 December 1990. Deaths related to tobacco smoking were excluded from the deceased mortality-matched referents, i.e. subjects were excluded if the following diseases were listed as underlying or contributing causes of death; cancer of the upper gastrointestinal organs, liver and biliary passages, pancreas, respiratory organs and urinary bladder, as well as ischemic heart disease, aortic aneurysm, bronchitis and emphysema, peptic ulcer, cirrhosis of the liver, and external causes. In total, information on 1042 cases and 2364 controls (including 1090 mortality-matched controls, of which 89% were deceased at time of data collection) was available for analysis. For the highest exposure category, a crude OR of 1.68 (95% CI: 1.23–2.30) was calculated. Adjustment for smoking, asbestos and some environmental factors had little effect on the risk estimate [OR = 1.63 (95% CI: 1.14–2.33)].

Comments on the study. The restriction of controls was justified by the study authors as follows; "we excluded smoking-related causes of death from the mortality-matched referent series, partly to obtain unbiased risk estimates of the effect of smoking [...]" (Gustavsson et al. 2000). The reference given in connection with this approach was a methodological study investigating the effect of excluding smoking-related deaths on the comparability of dead and living controls from a population-based case-control study on renal cancer with regard to their smoking habits (McLaughlin et al. 1985). With deceased controls more frequently having a history of smoking compared to living controls, the aim was to gain more similarity between the groups so as to be able to use deceased subjects as proper controls in studies on smoking as a risk factor. The principle of excluding control subjects with diagnoses that are caused by the traits under study has also been described by Rothman et al. (2008). Gustavsson et al. (2000), however, applied the principle without smoking being the exposure of interest. In their report, no risk estimate for the effect of smoking on lung cancer risk was given. Rather, the DE-lung cancer relationship was adjusted for smoking habits. As discussed earlier in the section on hospital-based case-control studies, excluding smoking-related diagnoses in such a situation widens the gap in the distribution of smoking habits between cases and controls. Taking into account that smoking is the most important confounder for the analysis of a DE–lung cancer relationship, it might be helpful to keep smoking-related diseases in the control group comparable to a matching procedure on smoking. We must assume that, by removing subjects with smoking-related diagnoses, the DE–lung cancer relationship is stronger biased than it would be the case if subjects with those diagnoses were included as controls.

In addition, the exclusion of smoking-related deaths disproportionately reduces the share of certain occupational groups in the control group. In light of the high smoking prevalence of DE-exposed workers, the proportion of smoking-related deaths will be comparatively high in this job group and, thus, the control group will be artificially deficient of DE-exposed workers. The result will be that the DE–lung cancer relationship is overestimated.

In the publication’s discussion, Gustavsson et al. (2000) also suggest that exclusion of smoking-related causes from the mortality-matched deceased controls prevented bias towards the null. They stated "We excluded smoking-related causes of death from the mortality-matched referent series, [...] because several of the risk estimates for the occupational exposure factors otherwise may have been biased towards the null, since several of these factors may increase the risk of death not only from lung cancer but also from other forms of cancer and cardiovascular diseases, in analogy with tobacco smoking". By this, the authors assume that DE might not only cause lung cancer, but also any other disease associated with tobacco smoking. As smoking is by far the strongest determinant of lung cancer, the extent of overestimating the DE–lung cancer relationship when excluding smoking-related deaths might be higher than the extent of underestimating risk when not excluding smoking-related deaths. Thus, we think that the risk estimate for the DE–lung cancer relationship given by Gustavsson et al. (2000) is an overestimation of the true association.

Summary. Residual confounding by smoking might be a simple explanation for the results of this study. The study cannot be considered as evidence for a causal relationship between DE and lung cancer.

Selection bias in population-based case-control studies

Overview of selection bias. The validity of population-based case-control studies to a large extent depends on the sampling procedure of control subjects. In order to compute an unbiased odds ratio, it is necessary to sample controls independently of exposure status from the source population. In principle, only complete participation of randomly selected controls ensures the absence of selection bias, provided that cases and controls stem from the same source population. In the presence of any incomplete participation of selected controls, participating and nonparticipating subjects must not differ from each other in their exposure (Rothman et al. 2008). However, participation of control subjects is seldom complete and, further to that, participation is frequently dependent on exposure status. For example, a survey on cardiovascular risk factors in the Swedish population aged between 25 and 74 has clearly shown a socio-economic gradient in the response rates (Strandhagen et al. 2010). The willingness to participate was significantly higher in individuals with university education [OR = 1.42 (95% CI: 1.29–1.56)]. The difference in terms of income is even greater. Moreover, among respondents, all health behaviors studied were significantly related to education.
<table>
<thead>
<tr>
<th>Reference (addition of papers for study description)</th>
<th>Study region and subjects</th>
<th>Recruitment of cases</th>
<th>Sampling of controls (PR: two control groups, of which one is matched to cases regarding vital status in 1990 with exclusion of smoking-related causes of death)</th>
<th>Response rate among controls (normal controls, 82% (mortality-matched controls)</th>
<th>Proxy respondents in cases/controls (93%/19% (normal controls, 89% (mortality-matched controls))</th>
<th>Adjustment for smoking (ever - never) 2</th>
<th>Adjustment for SES (cumulative)</th>
<th>Definition of exposure category, adjusted</th>
<th>OR (95% CI) for highest exposure category, adjusted</th>
<th>Trend with increasing exposure (n/a)</th>
<th>Considered in IARC's 2012 decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gustavsson et al. (2000)</td>
<td>Sweden, Stockholm County, men</td>
<td>Cancer registry, 1985–1990</td>
<td>PR; two control groups, of which one is matched to cases regarding vital status in 1990 with exclusion of smoking-related causes of death</td>
<td>88% (normal controls), 82% (mortality-matched controls)</td>
<td>93%/19% (normal controls, 89% (mortality-matched controls))</td>
<td>2</td>
<td>–</td>
<td>Cumulative</td>
<td>1.63 (1.14–2.33)</td>
<td>Possible</td>
<td>Yes</td>
</tr>
<tr>
<td>Parent et al. (2007) (Siemiatycki 1991)</td>
<td>Canada, Montreal area, men</td>
<td>All large hospitals, 1979–1985</td>
<td>PR (electoral roll, n = 375), RDD (n = 158); 1996–2001: PR (electoral rolls, n = 804)</td>
<td>72% (PR: 69.3%; RDD: 79.4%)</td>
<td>29.4%/12.6%</td>
<td>4</td>
<td>Family income</td>
<td>Ever – never</td>
<td>1.2 (0.8–1.8)</td>
<td>n.a.</td>
<td>Yes</td>
</tr>
<tr>
<td>Villeneuve et al. (2011) (Johnson 1998; Villeneuve et al. 1999)</td>
<td>Canada, 8 (of 10) provinces, men</td>
<td>Provincial cancer registries, 1994–1997</td>
<td>HC (5 provinces), RDD (2 provinces), property assessments database (1 province)</td>
<td>60.8%</td>
<td>0%</td>
<td>3</td>
<td>–</td>
<td>Cumulative</td>
<td>1.23 (0.9–1.7)</td>
<td>1.30 (1.3–2.6)</td>
<td>n.a.</td>
</tr>
<tr>
<td>Lechen et al. (1987)</td>
<td>New Mexico, USA, men</td>
<td>Tumor registry, 1980–1982</td>
<td>RDD, HC (additionally only for older subjects)</td>
<td>83%</td>
<td>50%/2%; not controlled for in analysis</td>
<td>2</td>
<td>–</td>
<td>Cumulative</td>
<td>1.12 (0.89–1.40)</td>
<td>p &lt; 0.07</td>
<td>n.a.</td>
</tr>
<tr>
<td>Richardi et al. (2006)</td>
<td>Italy, Turin</td>
<td>Hospitals, 1991–1992</td>
<td>PR</td>
<td>84.8%</td>
<td>0%</td>
<td>2</td>
<td>–</td>
<td>Education</td>
<td>1.26, p &lt; 0.05</td>
<td>1.35 (0.95–1.93)</td>
<td>n.a.</td>
</tr>
<tr>
<td>De Matteis et al. (2012) (Landi et al. 2008)</td>
<td>Italy, Lombardy region, men</td>
<td>13 hospitals with 80% coverage of all cases, 2002–2005</td>
<td>HC</td>
<td>72.4% or 57.2% (own computation) when relating participants to responding or</td>
<td>0%</td>
<td>4</td>
<td>–</td>
<td>Educational level</td>
<td>0.95 (0.63–1.45)</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Reference (additional papers for study description)</th>
<th>Study region and subjects</th>
<th>Recruitment of cases</th>
<th>Sampling of controls</th>
<th>Response rate among controls</th>
<th>Proxy respondents in cases/controls</th>
<th>Adjustment for smoking</th>
<th>Adjustment for SES</th>
<th>Definition of exposure</th>
<th>OR (95% CI) for highest exposure category, adjusted</th>
<th>Trend with increasing exposure</th>
<th>Considered in IARC’s 2012 decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrat et al. (2015) (Luce et al. 2011)</td>
<td>France, 10 regions (coverage 13% of total population), men</td>
<td>Cancer registries, 2001–2007</td>
<td>RDD</td>
<td>82%</td>
<td>Shortened direct or proxy interview: 8.9%/2.2%; not controlled for in analysis</td>
<td>4</td>
<td>(In subjects with no exposure to asbestos, the OR dropped from 1.62 to 1.47 when excluding subjects with highest education)</td>
<td>Cumulative</td>
<td>0.88 (0.65–1.19)</td>
<td>1.20 (1.0–1.44)</td>
<td>p = 0.36</td>
</tr>
<tr>
<td>Wild et al. (2012)</td>
<td>Northeastern France, men</td>
<td>Hospitals, 2006–2010</td>
<td>RDD</td>
<td>Not reported (possibly 66.4%)</td>
<td>Not reported (probably 0%)</td>
<td>4</td>
<td>Socio-economic class</td>
<td>Cumulative</td>
<td>1.20 (0.96–1.51)</td>
<td>1.66 (1.11–2.49)</td>
<td>p = 0.12</td>
</tr>
<tr>
<td>Menvielle et al. (2003)</td>
<td>France, New Caledonia (South Pacific), men</td>
<td>Cancer Registry, 1993–1995</td>
<td>PR (electoral rolls)</td>
<td>95.3%</td>
<td>(25.4%/&lt;1%)</td>
<td>3</td>
<td>–</td>
<td>Ever - never</td>
<td>0.8 (0.5–1.2)</td>
<td>n.a.</td>
<td>No</td>
</tr>
<tr>
<td>Tse et al. (2012)</td>
<td>China, Hong Kong, men</td>
<td>Largest oncology centre, 2004–2006</td>
<td>RDD</td>
<td>48%</td>
<td>0%</td>
<td>3</td>
<td>Education level</td>
<td>Cumulative</td>
<td>0.7 (no CI given)</td>
<td>2.10 (1.20–3.69)</td>
<td>no</td>
</tr>
<tr>
<td>Tse et al. (2011)</td>
<td>China, Hong Kong, men</td>
<td>Largest oncology centre, 2004–2006</td>
<td>RDD</td>
<td>Not reported (48% in study with full cohort, Tse et al. 2012)</td>
<td>0%</td>
<td>Analysis restricted to non-smokers</td>
<td>Education level</td>
<td>Duration</td>
<td>1.30 (0.68–2.49)</td>
<td>3.47 (1.08–11.14)</td>
<td>n.a.</td>
</tr>
</tbody>
</table>


*Smoking habits: 0 – no adjustment, 1 – adjustment for status (never, former, current), 2 – adjustment for status cross-classified by further information, 3 – packyears, 4 – packyears plus time since quitting.

No: trend not detected; (no): OR in highest category lower than in lower categories; possible – OR was highest in highest category; p-value given if reported; n.a.: not applicable

The number of interviewed controls (n = 894) divided by the given number of eligible controls (n = 1024) actually results in 86%; the given interview completion rate of 70% is probably derived from the number of eligible controls of 1294 reported in other publications about this study (Vallieres et al. 2012).

It is reported that 800 male controls initially declared willingness to participate; the study was run with 531 controls, however, which possibly would give a final interview rate of 66.4%
Based on the pooled analysis of case-control studies in Europe and Canada (Olsson et al. 2011) (Table 11), it was shown that the lung cancer risk in relation to DE is inversely correlated with the response rate among controls (Möhner 2012). Olsson and her coworkers showed that the pooled estimate for the DE-lung cancer relationship was somewhat biased due to one German study (the so-called AuT study). This study had the lowest response rate in controls among all pooled studies (response rate 41%) (Olsson et al. 2011, 2012). The main reason for this low response rate was probably the fact that the study base was part of a then ongoing study on lung cancer due to indoor radon exposure and the associated requirement to install a measuring instrument in the living room as well as in the bedroom for one year (Brüske-Hohlfeld et al. 1997, 1999).

Selection bias in the German study. We aimed to verify whether there was a difference in the prevalence of driver jobs between the control group in the German AuT case-control study and the general population. If there was a difference, the control group would not accurately represent the study source population with respect to an important job group for DE-exposure. In this situation, study results would be biased and fail to describe the true association between professional driving, DE and lung cancer. We also sought to estimate the variability of lung cancer risk estimates depending on how closely the control group resembled the source population.

Representative socio-economic data for the German population have been made available for scientific use since 2001 by the Research Data Centers of the Federal Statistical Office and the statistical offices of the federal states. We used data from the 1982 microcensus, a compulsory, representative, yearly re-run survey based on 1% of the population. Microcensus data from 1982 was limited to information on the population of West Germany, but for the first time included a three-digit job-code for each individual, referring to the German job classification system of 1975 (Statistisches Bundesamt 1975).

The detailed analysis on DE and lung cancer involving the West German part of the AuT study together with a second study in West Germany (HdA) shows a significantly elevated lung cancer risk for professional drivers (Table 13) (Brüske-Hohlfeld et al. 1997). The original dataset included 64 controls from West Germany that were economically active as professional drivers in 1982. In contrast, 100.37 drivers would be expected based on the microcensus 1982, taking into account age by 5-year age groups (Möhner 2016a). Therefore, the standardized employment ratio (SER), calculated in the same way as a SMR, is 0.64 (95% CI: 0.49–0.81).

A total of 426 cases and 234 controls in the western areas of the pooled German study were “ever employed” as professional drivers. In the analysis initially published, the figures provided were 412 and 226, respectively; however, a few individuals were categorized as non-exposed due to the data records showing that the vehicle was not powered by a diesel engine (Brüske-Hohlfeld et al. 1997, 1999). Assuming that the SER is 0.64 for all controls “ever employed” as professional drivers, the actual number of exposed controls should be 367 instead of 234. The impact of the SER on the risk estimates is given in Table 13. Strictly speaking, the SER is 0.64 only for controls employed as a driver in 1982. For drivers who changed jobs prior to 1982 it might be slightly higher. Therefore, an estimate of 0.7 seems realistic. However, the corresponding risk estimate, adjusted for smoking and asbestos, is no longer elevated under such an assumption (Table 13). This additional analysis demonstrates that population-based case-control studies may suffer considerably from certain bias introduced by a biased recruitment process of control subjects.

### Table 13. Lung cancer risk estimates in a German case-control study (Brüske-Hohlfeld et al. 1997) assuming different response rates among professional drivers in the control group.

<table>
<thead>
<tr>
<th>Work as a professional driver</th>
<th>Cases</th>
<th>Controls</th>
<th>OR</th>
<th>95% CI</th>
<th>OR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original distribution of control subjects (conditional logistic regression model)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>2263</td>
<td>2401</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>426</td>
<td>234</td>
<td>1.96</td>
<td>1.65–2.33</td>
<td></td>
</tr>
<tr>
<td>Original distribution of control subjects (unconditional logistic regression model)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>2263</td>
<td>2401</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>426</td>
<td>234</td>
<td>1.93</td>
<td>1.62–2.30</td>
<td></td>
</tr>
<tr>
<td>Model 1: SER = 0.64; nc = 133</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>2263</td>
<td>2268</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>426</td>
<td>367</td>
<td>1.16</td>
<td>1.00–1.36</td>
<td></td>
</tr>
<tr>
<td>Model 2: SER = 0.7; nc = 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>2263</td>
<td>2301</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>426</td>
<td>334</td>
<td>1.30</td>
<td>1.11–1.52</td>
<td></td>
</tr>
<tr>
<td>Model 3: SER = 0.8; nc = 58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>2263</td>
<td>2343</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>426</td>
<td>292</td>
<td>1.51</td>
<td>1.28–1.78</td>
<td></td>
</tr>
<tr>
<td>Model 4: SER = 0.9; nc = 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>2263</td>
<td>2375</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>426</td>
<td>260</td>
<td>1.72</td>
<td>1.45–2.03</td>
<td></td>
</tr>
</tbody>
</table>

SER: standardized employment ratio, i.e. the ratio of the prevalence of professional driving in the control group and the source population of the study; nc: Number of control subjects that change from unexposed to exposed status.

*Adjusted for smoking and asbestos.

bAssuming that the proportional risk reduction due to confounding factors is the same as in the unconditional logistic model based on the original data.
Selection bias in an Italian study. An analysis of non-response bias was also performed in the case-control study on lung cancer conducted in Turin (Richiardi et al. 2002). The researchers found that the socio-economic level of non-respondents was high in cases and low in controls. The prevalence odds ratio of “non-response” in control subjects was significantly elevated (OR ≈ 3) for the educational level “elementary school or less” in comparison to “high school and higher”. They concluded that “nonresponse, associated with socio-economic status, is an important potential source of bias in population-based case-control studies, which should always be considered and discussed”.

In a first analysis of the data on the relationship between DE exposure and lung cancer risk, the researchers reported an odds ratio of 1.04 (95% CI: 0.79–1.37) for any DE exposure, adjusted for smoking and exposure to list-A jobs (Richiardi et al. 2006). Further adjustment for educational level reduced the OR to 0.95 (95% CI: 0.72–1.26) (Table 12). The overall response rate among controls was 84.8%, about 80% in individuals with educational level elementary school or less and about 93% in individuals with high school education and higher (own calculations based on Richiardi et al. 2002 and Richiardi et al. 2006). The study in Turin, extended for the Venetian region, was later included into the pooled study by Olsson et al. (2011) (Table 11). The number of cases and controls was then almost doubled, but the overall response rate among controls decreased to 80%. In contrast, the risk estimate with respect to highest cumulative DE exposure without adjustment for education increased from 0.95 (95% CI: 0.63–1.45) (Richiardi et al. 2006) to 1.18 (95% CI: 0.89–1.57) (Olsson et al. 2011).

Selection bias in the Dutch study. A further example for the relationship between response rate and educational level is a case-control study nested in the Dutch sub-cohort of the European Prospective Investigation into Cancer and nutrition (EPIC)-study (Beulens et al. 2010). This study was also included in the pooled analysis by Olsson et al. (2011) (Table 11). The prospective Dutch cohort study comprised data from questionnaires and medical examinations of 22,769 participants aged 20–59 years. The overall response rate was 45% with older people more likely to respond than younger subjects (30% response among people aged 20–29 and 54% among people aged 50–59) (Beulens et al. 2010). Data on 64 lung cancer cases and 187 controls were included in the pooled analysis. The lifetime prevalence of occupational DE exposure among controls was 14.4%, significantly lower than the prevalence rate among all other study centers included in the pooled analysis. The second lowest value was 20.4% recorded in INCO-Romania. The HDa study in Germany, which adjoins the Netherlands, recorded exposure prevalence amongst controls of 57.6%. Taking into account that the age range of the Dutch sub-study was considerably lower than in the other sub-studies from the pooled analysis, one could speculate that the response rate was especially low among DE exposed persons that were still economically active. Indeed, the authors of the pooled study reported that the frequency of employment in jobs known to be associated with an increased lung cancer risk among controls was also the lowest in the Dutch sub-study (Olsson et al. 2011).

Control for selection bias in the data analysis. The adjustment for educational level, or any other indicator of socio-economic status, would assist in correcting the study results when differences in response rates between subjects exposed and not exposed to DE occur. However, of the separate population-based case-control studies reviewed (Table 12), four did not consider any control for socio-economic variables (Lerchen et al. 1987; Gustavsson et al. 2000; Menville et al. 2003; Villeneuve et al. 2011). Seven studies published main risk estimates adjusted for socio-economic variables (Brüské-Hohlfeld et al. 1999; Richiardi et al. 2006; Parent et al. 2007; Pintos et al. 2012; Tse et al. 2011; Tse & Yu 2012; Wild et al. 2012), while only two of these studies reported both unadjusted and adjusted results. In the Italian study adjustment for socio-economic status reduced the risk estimate by about 9% to an OR of <1, as reported above (Richiardi et al. 2006). A similar adjustment in the German study yielded an attenuation of almost 12% from an OR of 1.43 to 1.26 for lung cancer in relation to any occupational DE exposure (Brüské-Hohlfeld et al. 1999; Möhner et al. 1997). De Matteis et al. (2012) reported that adjusting for education did not have any impact on the risk estimates in their study. Matrat et al. (2015) did not adjust their main analysis for socio-economic information, but their sensitivity analysis restricted to subjects without asbestos exposure indicates confounding by educational level; the OR for all subjects was 1.62 and dropped to 1.47 when subjects were further restricted to those without a university or high school degree. Finally, additional adjustment for education in the pooled analysis of European and Canadian studies reduced the risk estimator in the highest quartile of exposed workers from 1.27 to 1.14 (95% CI: 1.03–1.26) (Straif et al. 2010) (Table 11). The final publication of this study, however, did not mention results adjusted for education (Olsson et al. 2011). In a later letter in response to Möhner (2012), Olsson and her colleagues reported the following: “Education was later dropped from the model, as it is not certain what attained education level reflects and if it is a real causal factor associated with lung cancer, after adjustment for other lifestyle factors such as smoking and occupational exposures to lung carcinogens, or that it is a correlate of DME exposure” (Olsson et al. 2012). Unfortunately, they did not report the adjusted estimate together with their argument why such adjustment might be critical in the first place.

Biased calculation of response rates. Last but not the least, the calculation of the response rates among the controls is worth discussing. In studies that used random digit dialing (RDD) for recruitment of controls, the response rate was mainly computed as the share of subjects finally taking part in the study among those that could be contacted by phone in the first place. The resulting response rates very likely overestimate the true rates, as eligible subjects that cannot be reached and thus remain unknown do usually not contribute to the calculation. This point is highly
relevant to DE exposure studies. Mobile workers such as long-haul truck drivers or heavy equipment operators, who often work away from home, are likely to be more difficult to contract compared with non-mobile workers. This at least holds true for landline telephone numbers, which were in use when the studies on DE exposure and lung cancer were conducted. Study personnel usually tried to contact a telephone number up to 5 or even 10 times at different times of the day and on different days of the week. However, even such a robust strategy cannot guarantee protection against a deficit of mobile workers among controls linked to RDD. Several of the reviewed studies employed RDD for the overall or a subset of the control group (Lerchen et al. 1987; Matrat et al. 2015; Parent et al. 2007; Pintos et al. 2012; Tse et al. 2011, 2012; Villeneuve et al. 2011; Wild et al. 2012) (Table 12). Parent et al. (2007) did not mention the use of RDD but in an earlier extensive study report it was noted that 158 of the 533 population-based controls had been recruited using RDD (Siemiatycki 1991; Siemiatycki et al. 1994). The same number of controls recruited using RDD was involved in the analysis by Pintos et al. (2012), yet the number of overall controls of 1427 was considerably larger in this extended study.

**Summary.** In population-based case-control studies, the selection of controls should optimally be based on a population registry or a comparable database. Only such data enables the researcher to investigate nonresponse accurately. Furthermore, investigators should reduce nonresponse bias by adjustment for educational level or similar socio-economic information. Four of the separate case-control studies on DE and lung cancer likely suffer from considerable selection bias arising from low response rates among population controls. In these studies, no additional adjustment for education or other factors describing the SES was performed. Such additional adjustment attenuates lung cancer risk estimates derived from case-control studies considerably, although it may not resolve this bias completely. Hence, selection bias together with unresolved confounding bias by smoking habits may fully account for the majority of the observed increased lung cancer risk in the reviewed studies. Consequently, the single studies and also the pooled study of European and Canadian study data should not be used for QRA.

**Discussion**

We performed a critical review of the literature that has recently been used by an IARC working group as the basis for the upgrade of DE to a Group 1 carcinogen. Our aim was to identify studies with sound methodology that could be used to evaluate whether a causal association between DE and lung cancer exists. A secondary goal was to select studies that could be used for a quantitative exposure–response assessment (QRA). The review brought to light a range of important topics that will be summarized in the subsequent discussion. This will be followed by our conclusions with regard to causality and risk assessment and recommendations for further research.

**Exposure estimation**

Miners working in underground worksites have the highest DE exposure intensity of the job groups analyzed in this review. Two large cohort studies assessed this group of workers, the DEMS conducted in the United States and the Potash miners study conducted in Germany. The mean REC exposure of production workers in the German potash mine study (154 μg/m³) was in the range of individual underground potash mines in the DEMS (122–219 μg/m³). The DE-exposure among underground workers in the DEMS was nearly two orders of magnitude higher than at typical surface workplaces with potential exposure to DE, including mining jobs at surface level.

Of the other studies reviewed, only three used quantitative exposure variables based on measurements of DE during work, two studies in the US trucking industry (Garshick et al. 2012; Steenland et al. 1998) and the study among DE-exposed workers in Finland (Guo et al. 2004). Thus, only a small fraction of studies reviewed could possibly be included in the QRA. Rather than quantitative exposure assessment, the other studies used years in exposed jobs for exposure–response analysis or simply compared ever-exposed subjects with the reference population. It is important to note that the high variability of exposure levels within a single job group is an important disadvantage of several studies. For example, DE exposure among truck drivers, the largest group among DE-exposed workers, depends on several factors such as engine power and maintenance intervals for the engine, weather conditions, vehicle density on the driving route, stop times for loading (when the engine is still running), and the ventilation practice of the driver in his cab. Exposure intensity among machine operators, particularly those working in the construction industry, and mechanics is dependent on the ventilation of the workplace.

Considering the high DE-intensity levels in underground mines, underground workplaces seem to be the most appropriate site for deriving robust estimates for a possible dose–response relationship between DE-exposure and lung cancer risk. However, any study in underground miners needs to carefully manage competing occupational risk factors such as exposure to respirable silica dust or radon exposure. Uncertainties due to backward extrapolation of exposure levels from measured values at a certain time point should also be taken into account, making the calculation of cumulative exposure levels prone to certain bias. Nevertheless, the coefficient of variation of data on DE-exposure seems to be more convenient in underground workplaces. This statement is supported by the replication of the analysis of the DEMS case-control study using alternative exposure metrics for DE, which showed the robustness of the results (Silverman et al. 2014; HEI 2015). However, the criticism voiced in this review on the analytical approach in the DEMS holds regardless of the exposure metric used and will again be highlighted in the sections below on overadjustment bias and model choice.

**Residual confounding by smoking**

Smoking is the strongest risk factor for lung cancer. As a comparison of 13 occupational categories demonstrated, the
prevalence of cigarette smoking was highest in employees in the transportation/material moving occupational category (Lee et al. 2004). This example shows that a thorough and efficient control for smoking is an essential prerequisite for the derivation of a valid estimate of DE-related lung cancer risk when subjects in exposed job groups are compared with subjects that might be quite different with respect to their smoking habits. The accuracy of adjustment for smoking depends on the precision and validity of available smoking data. This can be demonstrated using data of the German population-based case-control study on DE and lung cancer (Brüske-Hohlfeld et al. 1999). As shown in Table 14, the estimate for lung cancer describing the effect of DE-exposure is decreasing with increasing level of detail for the smoking information.

A further source of confounding bias is environmental tobacco smoke (ETS), originating from smoking drivers themselves or from their co-drivers. A recent survey in Australia showed that more than half of the long-haul truck drivers are exposed to ETS (Si et al. 2016). The exposure to respirable suspended particles PM_{2.5} originating from cigarette smoke may even be higher than that from the exhaust of a diesel heavy-duty truck, as recent studies demonstrated (Invernizzi et al. 2004; De Marco et al. 2016).

**Cohort studies**

Information about smoking in occupational cohort studies is usually scarce and imprecise. Moreover, the imprecision increases if the smoking information is gathered from next-of-kin or imputed from other sources. In studies with differences in the smoking habits between exposed and non-exposed group, residual confounding due to smoking is thus quite likely. As outlined in the comment on the US railroad worker study, observed differences in lung cancer risk could potentially be explained by differences of smoking habits. Moreover, in engineers, changes in the work environment over time entailed changes in smoking habits. The ideal situation of a cohort study in an occupational setting is thus to compare groups of workers defined by intensity of exposure, assuming that these groups are homogeneous with respect to other important factors such as age, birth cohort, education, lifestyle, and work environment. For the studies in miners, this set of parameters seems to hold true. As we have shown in our reanalysis of the published DEMS data, adjustment for smoking status and intensity did actually not influence the risk estimates for DE exposure. Therefore, the studies on miners seem to provide reliable results on a possible exposure–response relationship between DE and lung cancer risk among the available cohort studies.

The effects of birth cohorts present an additional bias with respect to smoking. Here again, the cohort from the US railroad retirement board is a good example. In this study, exposure accumulation and follow-up started in 1959 when subjects were between 40 and 64 years old. As a consequence, DE exposure was positively correlated with the year of birth. In general, a harmonization in respect to age between the exposed and unexposed subjects can be achieved, if study entry as well as the end of follow-up are adapted. However, this implies a strong reduction of person years in the given study. Even then an acceptable comparability may be not given. Table 8 shows that in the external reference population a large difference between the birth cohorts exist in terms of lung cancer mortality. The age-specific share of lung cancer in total cancer deaths in the youngest birth cohort is almost twice as high as in the oldest birth cohort [own calculations based on US mortality statistics (Ribicoff & Terry 1961; CDC-WONDER 2009)]. This implies that smoking in younger subjects plays a much stronger role for the lung cancer mortality than in older subjects. In turn, confounding by smoking might then disproportionally affect younger birth cohorts with longer exposure experience. However, this difference is not reflected by the smoking data imputed for use in the cohort analysis of the railroad worker study.

**Case-control studies**

The availability of smoking data is generally better in case-control studies as they offer the opportunity to collect information about possible confounding factors from the participants themselves or from their relatives. However, the issue of potential residual confounding by smoking cannot be completely excluded as an assessment of smoking habits is complex. As previously discussed, most of the hospital-based case-control studies and one population-based study (Gustavsson et al. 2000) additionally excluded subjects with smoking-related diseases or causes of death from their controls. This is likely to have exaggerated the confounding problem. Lung cancer risk estimates may have been overestimated by artificially making the control group more different from the case group with regard to the most important risk factor for the disease, smoking.

**Selection bias in case-control studies**

Selection bias among population-based controls in the case-control studies is likely to have further biased results.

<table>
<thead>
<tr>
<th>Smoking related variables</th>
<th>Other variables</th>
<th>OR</th>
<th>95%CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>–</td>
<td>–</td>
<td>1.91</td>
<td>1.68–2.18</td>
</tr>
<tr>
<td>ln(PY), ex-smoking(3), pipe/cigar(1)</td>
<td>–</td>
<td>1.47</td>
<td>1.26–1.70</td>
</tr>
<tr>
<td>ln(PY), ex-smoking(3), pipe/cigar(1)</td>
<td>Asbestos(1)</td>
<td>1.43</td>
<td>1.24–1.67</td>
</tr>
<tr>
<td>ln(PY), ex-smoking(3), pipe/cigar(1)</td>
<td>Asbestos(1), SES(3)</td>
<td>1.27</td>
<td>1.09–1.48</td>
</tr>
<tr>
<td>Current, former, never (Garshick et al. 2008)</td>
<td>–</td>
<td>1.63</td>
<td>1.41–1.88</td>
</tr>
<tr>
<td>Never, 1–49 PY, 50+ PY (Garshick et al. 1987a, 1987b)</td>
<td>–</td>
<td>1.64</td>
<td>1.43–1.88</td>
</tr>
</tbody>
</table>

PY: pack years; SES: socio-economic status; Number of parameters in parentheses.

*Comparison of ever DE-exposed vs. never exposed.*
Generally speaking, the aim is to select controls in a way that represents a blueprint of the source population (i.e. the base population that gives rise to the cases). However, for DE exposure, this aim is difficult to achieve using traditional sampling strategies as the response rate among control subjects is correlated with SES. The result is an underrepresentation of blue-collar workers and, thus, an underestimation of the exposure prevalence in the control group. The controls are therefore likely to differ from the source population in terms of lifestyle habits and working conditions. Additionally, the selection of highly-mobile subjects, such as truck drivers, is often difficult as they are often away from home. This job group is also likely to be underrepresented in the control group. The resulting bias in the selection of control groups may be responsible for the increased risks of lung cancer in exposed study groups and hinders the assessment of observed risk differences. Individual matching via detailed information on smoking habits or SES might reduce residual confounding. However, this is not usually performed. Case-control studies nested within large cohorts of workers with similar work tasks and similar work environment provide generally better opportunities to derive reliable risk estimates.

**Overadjustment bias**

**Adjustment for a supposed healthy-worker effect**
The healthy-worker effect is always an issue in occupational epidemiology. Several strategies have been developed in order to deal with this kind of bias (Arrighi & Hertz-Picciotto 1994; Arrighi & Hertz-Picciotto 1995; Steenland et al. 1996; Li & Sung 1999; Baillargeon 2001; Richardson et al. 2004; Applebaum et al. 2007, 2011; Chevrier et al. 2012; Joffe 2012; Naimi et al. 2013; Picciotto et al. 2013; Buckley et al. 2015), but as recently formulated: the HWSE is a still-evolving concept (Picciotto & Hertz-Picciotto 2015). Some of the strategies referenced above do not take into account the primary aim of occupational epidemiology, which is to investigate the relationship between exposure and the first diagnosis (incidence) of a disease. Mortality serves merely as a surrogate for incidence. Hence, an appropriate approach when considering a possible HWSE should take into account the lethality rate of the disease of interest. A complete discussion of this phenomenon is beyond the scope of this paper. We will therefore discuss only the aspects we believe are important for the cohort studies on DE exposure.

Lung cancer has a high lethality rate (Compton et al. 2012). Therefore, exposure lagging by five years should be sufficient to prevent HWSE. However, two of the large cohort analyses have been adjusted for years worked and years off work in order to control for a supposed HWSE. This was true of the railroad worker cohort of the US railroad retirement board (Garshick et al. 2004, 2006; Laden et al. 2006) and the cohort of employees in the US unionized trucking industry (Garshick et al. 2008, 2012). Here, the study authors did not explain why the total years employed in the corresponding industries should be treated in the analysis as a confounder. In our opinion, the consideration of years worked as a time-dependent variable in the Cox models leads to biased risk estimates for the following two reasons. First, this variable does not cover the complete former time of employment without DE-exposure leading to a differential bias. Second, the years worked are strongly correlated with cumulative exposure, which induces an overadjustment bias. A less problematic consideration of the employment duration could be achieved if the years worked without DE-exposure would be used instead of total years worked in the corresponding industry. The years worked with DE exposure are already incorporated into the cumulative exposure variable.

With regard to the adjustment for years off work the use of mortality data in the absence of incidence data needs to be considered again. Most lung cancer patients, as well as workers with other serious illnesses, will not remain employed in their jobs following their diagnoses. Consequently, the SMR is strongly increased in the first year after a premature termination of employment. The relationship between time since premature termination of employment and mortality is related to the lethality rate of the disease under investigation. In the case of lung cancer the lethality rate is high. Thus, the SMR in the first year after a premature termination of the employment is also considerably elevated. However, this observation alone cannot justify an adjustment for years off work.

In conclusion, we do not see the necessity to adjust for total years worked or years off work in the analyses on DE and lung cancer. Moreover, we suspect bias in conjunction with such an approach. In this case, a reasonable approach to adjust for HWSE is exposure lagging, as was performed in many studies on this topic. A lag time of 5 years seems to be sufficient to prevent a HWSE.

**Adjustment for work location in the DEMS**

When the variable work location was included in the statistical model of the DEMS data, a relationship between DE and lung cancer was observed. This adjustment is questionable as the work location variable is tightly correlated with the exposure under question and any subsequent adjustment procedures lead to overadjustment bias and the suggestion of a strong exposure–response effect. Although the positive relationship in the subgroup of ever-underground workers from the cohort analysis first suggested such a relationship, this observation may actually be the result of a healthy-worker survivor bias instead.

**Model choice**

The internal analysis of cohort studies is in most cases performed by Cox or Poisson regression. Logistic regression is a standard approach for case-control studies. There are many other ways to model the relationship between the exposure and the outcome, particularly if a set of possible confounding factors is included. When the ratio between the number of observations/cases and number of parameters to be estimated falls below an acceptable range, the precision of the parameter estimates decreases rapidly and “sparse-data” bias can result (Greenland et al. 2000). In this case methods are required to balance the model bias and the estimation error.
AIC and BIC are both methods of assessing model fit and penalize free parameters in an effort to combat overfitting. BIC tends to penalize models more for free parameters than does AIC. When applied to the DEMS, both measures argued for the use of a simple model, including main effects for the exposure variable (either work location or cumulative REC) and smoking status with or without smoking intensity. In contrast, the combination of work location cross-classified by smoking status and smoking intensity as used by the DEMS authors was identified as the inappropriate option.

Cross-validation can also be useful in selecting the appropriate statistical model for an analysis (Arlot & Celisse 2010). A suitable cross-validation approach for the DEMS could be to take the data for seven of the eight mines to generate the risk estimates for various models. This estimate is then used for prediction in the eighth mine and to calculate the precision of the fit. This procedure should be repeated by rotating the mines in the analysis until all mines have been used for testing the model fit. The model with the best fit, summarized over all eight runs, is then the recommended model to use for the pooled analysis.

**Synopsis of the results**

Of the studies reviewed, only the cohort-based studies among nonmetal miners are suitable for derivation of valid quantitative lung cancer risk estimates in individuals occupationally exposed to DE. The DEMS has the most informative database. However, the unusual adjustment of REC exposure for smoking cross-classified with work location led to strong overadjustment bias. The study authors concluded that a causal relationship exists between DE-exposure and lung cancer risk. An alternative hypothesis for the observed relationship would be the presence of a healthy worker effect. Our rough reanalysis of the DEMS yielded only a very flat non-significant increase of lung cancer risk with increasing DE exposure, similar to the results of the reanalysis of the German Potash Miners Cohort study.

The results of the studies in the unionized US trucking industry cannot be used because of biased estimates. The other study with quantitative exposure data – the Finnish registry-based study on DE exposed workers (Guo et al. 2004) – adds some evidence against a DE–lung cancer relationship, although this study also has noticeable limitations.

The other cohort studies reviewed suffer from significant limitations with respect to exposure assessment. For some studies, the assessment of exposure via the quantification of exposure by years of work since 1959 in a certain job introduces further bias. An exposure classification by cumulative DE exposure is, in fact, the same as a classification by birth cohorts. Hence, birth cohort effects with respect to confounding factors, especially to smoking, and selection processes may substantially bias the results. The large-scale cohort study in the railroad industry (Garshick et al. 2006) provides further evidence against a link between DE and lung cancer risk as the study results can be easily explained by factors such as smoking.

Population-based case-control studies are limited by the fact that response rates among controls are correlated with socio-economic parameters such as education. A response rate below average in DE exposed workers entails an overestimation of the corresponding lung cancer risk. Unfortunately, for only a few studies data were available, allowing for some verification of this type of correlation. Crude calculations based on these data indicate that the elevated risk estimates were at least greatly diminished after correcting for heterogeneity in the response rates by socio-economic status. Hospital-based studies were likely biased by excluding subjects with smoking-related diagnoses from the control groups. Therefore, case-control studies can add only little evidence for the DE–lung cancer hypothesis.

**Conclusions and recommendations for further research**

**Conclusions with respect to causality**

This critique of the methodological issues seen in occupational epidemiological studies strongly suggests that there is currently insufficient evidence to confirm the hypothesis of a causal link between DE exposure and lung cancer risk. The body of evidence could be enhanced by a re-analysis of the most informative study base, the DEMS. In order to appraise the researchers’ post-hoc hypothesis (Silverman & Attfield 2012) correctly, the case-control approach could be reduced to surface-only and underground-only workers. A second suitable approach would be to include two separate exposure variables for DE; the first for exposure on surface and the second for exposure underground.

Verification of the alternative HWE-hypothesis (Möhrner 2016b) should be based on a classification of workers by work location at study entry. Beside smoking status, former employment in other underground mines, former employment in other jobs classified as list-A-jobs (Ahrens & Merletti 1998; Mirabelli et al. 2001), duration of mining outside the study mines, and the year of hire in an underground job in a “study mine” should be taken into account. Unfortunately, the authors of the DEMS seem to ignore this alternative hypothesis (Silverman et al. 2016). However, we would like to reiterate that the classification into surface-only and ever-underground workers is not an appropriate classification for the verification of their post-hoc hypothesis.

**Conclusions with respect to risk assessment**

The DEMS results were already used for QRA (Vermeulen et al. 2014; MacCalman et al. 2015; Neophytou et al. 2016). However, risk was overestimated in these calculations, due to the methodological issues in the DEMS described above. There are only the two studies among nonmetal miners, which are suitable for QRA. Of these, the DEMS needs to be re-analyzed in order to exclude overadjustment bias and to select an appropriate analytical model.

**Recommendations for further research**

Taking into account the huge difference in exposure intensity between underground and surface jobs, as well as the
uncertainties concerning the exposure assessment for almost all DE exposed jobs on surface, epidemiological data on miners should ideally be used to assess a possible exposure–response relationship between DE and lung cancer risk. A pooled analysis of the two studies among nonmetal miners could further improve the validity of the results. Moreover, we recommend the use of a uniform lag time in all studies on DE and lung cancer. All future studies should at least include results with a 5-year lag-time.

It should be noted, that measurements in civil engineering have revealed that workplaces exist in this field where DE exposure is much higher in comparison with the studies in underground mines discussed above. Exposure concentrations of ~1.5 mg/m³ REC were measured in deep excavations or trenches operating diesel-powered rammers or vibration plates (Ziegler et al. 2014). It is also conceivable that some underground workplaces such as tunnel constructions are prone to DE concentrations in that range. Epidemiological data on workers exposed to such intensity levels of DE exposure are missing. It is known from animal studies that the lung cancer risk in rats can be described by a threshold model (Valberg & Crouch 1999). Following these calculations, the threshold can be expected between 160 and 600 µg/m³ average continuous lifetime exposure. Further epidemiological research is urgently needed that includes high exposed workplaces to verify such a threshold model for humans.

**Recommendation for a threshold value**

Our review shows that most studies cannot add evidence for a causal link between DE exposure and lung cancer risk. Therefore, a reliable derivation of a quantitative exposure–response relationship is not possible at present. In view of the results from animal studies and the fact that a threshold model cannot be ruled out, a conservative lower bound for a possible threshold value should be determined. Such a value could be derived from the German cohort study among potash miners. An upper bound for the cumulative exposure of 2.5 mg/m²-years REC seems to be sufficient to prevent a detectable increase of lung cancer risk. This value corresponds to an average annual value of 50 µg/m³ REC assuming a working life of 45 years. Unless a re-analysis of the DEMS yields considerably higher values, this value could be recommended as threshold value for occupational safety.

**Acknowledgements**

The authors gratefully acknowledge the careful review provided by the anonymous reviewers. We also thank Dr. Norbert Kersten, Dr. Dagmar Pattloch and Anne Pohrt for their helpful comments on an earlier version of this manuscript. Finally, we thank Dr. Clare Barnett for language editing.

**Declaration of interest**

The employment affiliation of the authors is as shown on the first page. The Federal Institute for Occupational Safety and Health (BAuA) is a German Federal Departmental Research Institute. As a federal authority, it is directly responsible to the Federal Ministry of Labor and Social Affairs (BMAS). It conducts research and development in the field of safety and health at work, promotes the transfer of knowledge into practice, advises policymakers and performs sovereign functions (www.baua.de).

As part of their work activities, the authors act as advisors in several committees in Germany that discuss and provide policy advice on occupational health issues. MM is an advisor for the Permanent Senate Commission for the Investigation of Health Hazards of Chemical Compounds in the Work Area (MAK Commission) which proposes maximum workplace concentrations (MAK values) for volatile chemicals and dusts, biological tolerance values (BAT values), biological guide values (BLW), biological reference values for workplace substances (BAR) and analytical methods for substances in the air and biological material. Substances, which are carcinogenic, germ cell mutagenic, sensitizing or absorbed percutaneously or which pose a risk during pregnancy are classified accordingly (www.dfg.de/en/dfg_profile/statutory_bodies/senate/health_hazards/index.html). MM further is an advisor for the subcommittee “Risk Assessment” (UA III) of the Commission on Hazardous Substances (AGS), which counsels the BMAS on work safety with regard to hazardous substances, including classification and labeling. Evaluation of the risk associated with occupational exposure to diesel motor emissions is a current issue in both committees.

Furthermore, MM and AW provide epidemiologic advice for the Medical Expert Advisory Board Occupational Diseases, which counsels the BMAS on the scientific prerequisites with regard to new legal occupational diseases.

The authors take full responsibility for the analyses, interpretations and conclusions in the present review. The presented ideas and opinions are exclusively those of the authors and do not necessarily reflect those of the mentioned institutions.

**References**


Morfeld P. 2012. Diesel exhaust in miners study: how to understand the findings? J Occup Med Toxicol. 7:10


StataCorp. 2015. Stata: release 14. Statistical software. College Station, TX: StataCorp LP.


Wild P, Gonzalez M, Bourgkard E, Courouble N, Clement-Duchene C, Martinet Y, Fargette J, Paris C. 2012. Occupational risk factors have to...


