AN APPLICATION OF SEMI-BAYES MODELING TO A STUDY OF THE OCCUPATIONAL ETIOLOGY OF LUNG CANCER

FRANCO G MOMOLI

Department of Epidemiology and Biostatistics
McGill University, Montreal
December 2005

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements of the degree of
Doctorate of Philosophy

Your file Votre référence
ISBN: 978-0-494-25214-7
Our file Notre référence
ISBN: 978-0-494-25214-7

Doctorate of Philosophy (2005) McGill University
Montreal, QC

Title:	An application of semi-Bayes modeling to a study of the occupational etiology of lung cancer
Author:	Franco G Momoli
Supervisors:	Drs. Jack Siemiatycki & Michal Abrahamowicz
Co-supervision:	Dr. Marie-Élise Parent
Number of Pages:	xiv, 305

Table of contents

List of figures

List of tables

List of appendices

Statement of support

During the time spent on the present thesis, I was supported by a Doctoral Research Award from the Canadian Institutes of Health Research. Personal funding was also provided from the Canadian Cancer Society through a grant by the National Cancer Institute of Canada and the Program of Research in Environmental Etiology of Cancer.

Acknowledgments

How fortunate I have been to see through the eyes of Michal Abrahamowicz, Marie-Élise Parent, and Jack Siemiatycki, to share in their insights and in their humour. I offer my sincerest thank-you to Jack for supporting me all these years, and especially for his patience during the times that life interrupted science. I can only say that I'm grateful for all those conversations on epidemiologic theory, for lending me the mini-van on occasion, and for providing me the right atmosphere and team to pursue interesting ideas (that CSEB newsletter got it right in naming him a folk hero). Marie-Élise sat patiently through our daily 21 questions, jumping from occupational hygiene to mathematical statistics to parenthood, while tossing in the occasional lesson on how I misuse English words, such as drank and drunk. I believe I would not have made it, were it not for her constant nudging. A conversation with Michal always left me with the impression that I had spoken of important matters intelligently – but I know this was his talent. I shouldn't understate how many times the clarity of his thought bailed me out of the confusion of my own, and how many times his enthusiasm for the ideas in the thesis revitalized me.

A special thanks to Ramzan Lakhani, Benoit Latreille, and Louise Nadon, for the generous amount of time they spent working out those "baskets." And thank-you to the rest of the teams at IAF and CHUM, especially Sally Campbell, Marie Désy, and Sandra Woods-Lecourt, for all the little things that helped me along the way.

Robert Platt, Lawrence Joseph, and Sander Greenland offered answers to (what seemed to me) perplexing questions concerning empirical Bayes models. I also had an invaluable correspondence with Anneclaire de Roos, who often served as my sounding board for many aspects of the semi-Bayes models. My thanks to Marie-Claude Rousseau, who may not have realized it, but helped to clarify several epidemiologic issues that nagged at me; Andrea Benedetti, for being there to discuss all those conceptual and computational issues, many of which I wish in the end had made it into my thesis; Karen Leffondré, for advice on some nasty statistical issues, mostly having to do with interpreting her 'special smoking variables'; Jérôme Asselin, for translating my abstract to French; Raman Agnihotram, for interesting discussions; Michel Camus, for suggestions about the title of the thesis; Dan Krewski, for kindly offering me a desk in Ottawa to finish up my writing;

and Javier Pintos, for having all the answers to all the questions ... but mostly for keeping his operatic singing to a minimum.

Thank-you to Emma's parents, who have been so good to me during this crazy time -- among other things, they provided me a place to stay during all those trips to Montreal. And thanks to Andrea Benedetti and Tom Koutsavlis, for putting me up several times, disparaging remarks about Star Wars notwithstanding.

(... and would you think me crazy for thanking Daniel Lanois and Peter Gabriel for their respective albums, 'Shine' and 'Up'? They calmed my soul when it was in need of calming ...)

My mother and father have supported me through a seemingly never-ending education. Though I could always hear the worry in their voices about whatever was troubling me on any given week, I hope they know that I have always drawn confidence and strength from their love. It is *their* outlook on life that I would hope to pass on to my children. Ava and little Jack have brought me so much joy these last years. I find it hard to fathom how truly wondrous they are. I only hope they are too young to be jaded by how many decades my schooling has taken me, and that it won't turn them away from the joys of education entirely. And my wife, Emma, was tolerant of what could only have been incessant babbling on P-values and Bayesian statistics. She held down the fort in the interminable dying days of my thesis writing, and she provided me with backbone when I thought all was going wrong. To her, I express my love, and I hope that with our student days finally behind us, we can start to live the good life.

I leave off with a Yiddish saying that is perhaps the silent motto of every thesis ever written: Troubles overcome are good to tell.

Statement of originality

The present thesis is based on data collected in a study designed by Dr. Jack Siemiatycki. The purpose of this doctoral work was to evaluate several hundred occupational chemicals for their effects on lung cancer, using various modeling techniques. While several of these chemicals have been assessed in other studies by other researchers, and some screening analyses of the data concerning lung cancer from the present study have been previously published, my own work aimed at more comprehensive and advanced analyses of the effects of the chemicals.

An important feature of this work addressed mutual confounding among a large and complex set of exposures. This is rarely given sufficient consideration in occupational cancer studies, because relevant information about confounding exposures is often limited. Another important contribution of this work was the application of an established but rarely used analytical approach, namely semi-Bayes modeling, which was designed to handle problems that arise in the assessment of multiple exposures. By applying this analytical technique, I could fit a much larger regression model than is normally possible, and also quantitatively incorporate expert opinions about chemical properties.

I established the framework for formulating the type of expert knowledge required for the semi-Bayes model, and under my supervision the chemical properties listed in Table 6-7 were designed by the following occupational hygienists: Dr. Louise Nadon, Benoit Latreille, and Ramzan Lakhani.

To my knowledge, this is one of the largest simultaneous assessments of occupational chemicals and lung cancer to date. The results provided evidence on the risk of lung cancer in relation to many exposures for which little is known about their carcinogenic effects.

Abstract

The occupational environment has been a fruitful source of research on causes of cancer. Analyses in studies of occupational risk factors for cancer can experience problems if an attempt is made to model large numbers of exposures, some of which may be highly correlated. Typical analyses of such studies focus on one chemical at a time, but this may not adequately deal with mutual confounding. Based on a large study in Montreal, the objective of this thesis was twofold: to assess several occupational chemicals for their etiologic role in lung cancer, and to explore the use of semi-Bayes modeling to simultaneously estimate the effects of many chemicals at a time. METHODS: Data came from a multiple-cancer case-control study of exposures in the work place. The study was comprised of 857 cases of lung cancer and 2172 controls consisting of patients with other types of cancer diagnosed from 1979 to 1985. Detailed occupational histories were collected and occupational hygienists translated these into exposure histories for 231 chemicals. All chemicals were analysed with conventional modeling strategies of both single and multiple parameter models. Of the 231 chemicals, 184 were singled out for analysis in a single large semi-Bayes model, which is a variant of classical empirical Bayes. This analysis is a fairly novel method suited to estimating large numbers of parameters in the face of sparse data. For the Bayesian portion of this model, chemicals were grouped by shared chemical and physical properties, based on the belief that these shared properties would imply similar effects on the risk of lung cancer. RESULTS: Estimates for all 231 chemicals were derived under the various modeling strategies. For most chemicals, estimates changed little across these analytic approaches, though some differences were apparent. Of the 231 chemicals assessed, 53 were earmarked as requiring further evaluation and underwent additional analyses. DISCUSSION: While semi-Bayes models have been shown previously to offer improved estimation over conventional analyses, the gains in using semi-Bayes models in the present study were less clear. Effort put into some portions of the Bayesian modeling did not materially influence the results. A number of chemicals were earmarked as potential lung carcinogens.

Résumé

L'environnement professionnel offre de vastes opportunités de recherche sur les causes du cancer. Dans de telles études, des problèmes sont observés lorsque est tentée la modélisation d'un grand nombre d'expositions dont certaines peuvent être fortement corrélées. Ainsi, les analyses typiques ne ciblent qu'un produit chimique à la fois, sans tenir compte de possibles effets de confusion engendrés par d'autres substances. Basée sur une large étude montréalaise, cette thèse comporte deux objectifs principaux: évaluer le rôle étiologique de plusieurs produits chimiques sur le cancer du poumon; explorer les modèles semi-bayésiens pour l'estimation simultanée des effets de plusieurs produits chimiques. MÉTHODES: Les données proviennent d'une étude cas-témoins portant sur des cancers multiples en relation avec l'exposition en milieu professionnel, à laquelle ont participé 857 cas de cancer du poumon et 2172 témoins ayant d'autres types de cancer. L'historique professionnel des patients fut récolté et traduit en expositions à 231 produits chimiques par des hygiénistes du travail. Tous les produits furent analysés par les méthodes conventionnelles avec modèles simples et multiples. Des 231 produits, 184 furent retenus pour une large analyse multiple avec un modèle semi-bayésien, une variante du modèle empirique de Bayes. Cette récente technique d'analyse s'avère efficace pour estimer un grand nombre de paramètres avec des données dont le nombre est limité. Pour la portion bayésienne de ce modèle, les substances furent regroupées par propriétés chimiques et physiques, sous l'hypothèse qu'à ces propriétés partagées seraient reliés des effets similaires sur les risques de cancer du poumon. RÉSULTATS: Les estimations de risque pour l'ensemble des 231 produits chimiques furent obtenues sous les diverses stratégies analytiques. Pour la majorité des produits, les deux méthodes donnaient des résultats comparables, avec quelques exceptions. Des 231 produits étudés, 53 furent retenus pour analyses plus approfondies. DISCUSSION: Bien qu'il ait été démontré par le passé que les modèles semi-bayésiens résultaient en une amélioration des estimations comparativement aux analyses conventionnelles, les gains découlant de l'utilisation des modèles semi-bayésiens dans cette étude ne sont pas aussi évidents. Les efforts investis dans l'application de ces derniers n'ont pas influencé substantiellement les résultats. Un certain nombre de produits chimiques ont été retenus comme carcinogènes pulmonaires potentiels.

Terminology

The phrases 'occupational etiology' and 'occupational cancer' refer to human risks attributable to hazards in the workplace. These hazards can relate to general work conditions and environments or, as in the present thesis, chemical exposures common to, but not necessarily limited to, the workplace.

As a matter of style, capital roman letters are used to represent parameters (in lieu of their Greek equivalents), small roman letters represent empirical estimates of the parameters, and bold letters signify vectors or matrices of parameters or estimates. The following acronyms are occasionally used: RR for incidence density rate ratio, SMR for standardized mortality ratio, PMR for proportionate mortality ratio, OR for odds ratio, EB for empirical Bayes, SB for semi-Bayes, and MLE for maximum likelihood estimate. Among the chemicals, the following acronyms were used: PAH for polycyclic aromatic hydrocarbons, MAH for monocyclic aromatic hydrocarbons, and B(a)P for benzo(a)pyrene.

The expression 'confidence limits' (CL) is used to generically describe even the interval estimates from the empirical Bayesian and semi-Bayes analyses, though the use of the word confidence, as it implies a Frequentist method, is formally incorrect. The terminology of empirical Bayesian modeling varies greatly throughout the published literature. Such approaches are equivalent to, or special cases of, hierarchical or multilevel models. The terminology used in the thesis will refer to the model containing the parameters of interest (for the occupational exposures) as the first-level model or the maximum likelihood model, and the empirical Bayesian prior information (*about* the first-level parameters of interest) as the second-level model.

Reports of confidence limits throughout the introduction and literature review can be assumed to be 95% intervals, unless otherwise specified.

1 Introduction

Dangerous substances and conditions seem to have always existed in the workplace, though they were not necessarily a source of concern for public health. 'Unhealthy' work environments were occasionally exploited as forms of punishment for slaves and convicts (Hunter, 1975), but it has been mostly ignorance, negligence, and maximization of profits that have maintained the unhygienic workplace. Consideration of the occupational causes of cancer has been in evidence for some time. A 1556 text appears to describe symptoms of respiratory disease, including lung cancer, in miners (Agricola, tr1950). In the 18th century, Bernardino Ramazzini (tr1964) wrote that physicians, in addition to noting signs and symptoms of illness, should inquire about occupation. But actual investigation into occupational causes likely dates back to Percival Pott (1775), a surgeon at Saint Bartholomew's Hospital in London. Through his own clinical observations, he described scrotal cancer in chimney sweeps, attributing the cancer to the cumulative effect of deposits of soot.

Need for detailed occupational information was noted in 1851 when the Registrar General's Office in England found it impossible to determine relative mortality of silk, cotton, linen and woollen manufacturers, since most were registered simply as weavers (Farr, 1975). The history of occupational studies in epidemiology is such that most research has been conducted by relating cancer mortality to occupational or industrial titles. Relatively few studies addressed the onset of cancer and the role of exposure to particular chemicals. With so many chemicals introduced into the workplace annually, it has been recognized that there is a need to assess the health effects of exposure to chemicals whose effects have not yet been investigated (Siemiatycki et al., 1981). Past research on occupational cancer may well have only uncovered the "tip of the iceberg" of occupational carcinogens.

Of all cancers, lung cancer has been associated with the most occupational causes discovered to date (Siemiatycki et al., 2004). That it is also a leading cause of death in much of the world, and that it has such a poor survival rate, underlines the importance of knowledge leading to primary prevention, such as that provided by occupational cancer research.

What follows is a brief overview of the descriptive epidemiology and etiology of lung cancer.

1.1 Descriptive epidemiology of lung cancer

Mortality rates and incidence rates of lung cancer have been similar because of the high case-fatality rate of the disease, and so routine collection of vital statistics has provided a long historical record of the occurrence of lung cancer (Alberg and Samet, 2003). Lung cancer was rare until about 1930, when a sharp rise in incidence eventually led it to becoming the leading cause of death from cancer among men. This is largely consistent with the sharp rise in cigarette smoking among men through the first half of the last century (Burns, 1994). Women's smoking habits increased several decades after men's and their lung cancer rates rose accordingly later than men's did. Though men have tended to have a twofold higher incidence of lung cancer than women (Miller, 1983), Lubin and Blot (1984) showed that differences between men and women were not entirely due to smoking, suggesting sex-dependent factors may have a role in lung cancer etiology. But for the most part, the rise and decline of rates of incidence of lung cancer parallel trends of tobacco use (Wingo et al., 1999), and many of the observed temporal, geographic, and demographic variations mentioned below are partly or wholly accounted for by patterns of cigarette use.

Marked variation in lung cancer rates occurs geographically (Mason, 1994), though not all of this variation is thought to be due to smoking practices. For example, Western Scotland had a higher rate of lung cancer compared to the U.S., at all levels of cigarette smoking (Hanai et al., 1988), suggesting other factors were at play, perhaps low consumption of fresh fruit and vegetables (Gillis et al., 1988). Incidence rates tend to be higher in developed countries, even taking into account differences in diagnostic practices (Pisani et al., 1999). Urban areas in most parts of the world were also found to have high rates, leading to the hypothesis that air pollution may contribute to lung cancer, though the urban-rural difference has been attributed by Doll and Peto (1981) to patterns of cigarette smoking. High rates of lung cancer in U.S. coastal regions led to studies assessing the ship building industry and asbestos exposure (Blot et al., 1978).

Rates of lung cancer have also been found to vary with race, ethnicity, social status, and education. For example, black men appear to have 50% higher rates than white men (Blot and Fraumeni, Jr., 1996), and all these rates tend to be higher than those of Hispanics, Native Americans, and Asians. Income, education, and social economic status have been found to be inversely related to lung cancer (Fraumeni, Jr. and Blot, 1982; Mao et al., 2001). Jewish men have been found to have very low rates of lung cancer, likely due to less tobacco smoking, but the reduced risk does not appear to exist with Jewish women (Horowitz and Enterline, 1970).

Several histologically distinct types of lung cancer exist, arising almost entirely from epithelial tissue (Beadsmoore and Screaton, 2003). The most common types are squamous cell carcinomas, adenocarcinomas, and small cell carcinomas. These categories, while formally defined, are often not reliably differentiated by pathologists (Ives et al., 1983). Further, tumours can transform over time from one type to another, and a tumour can manifest as combinations of types. Each stage of tumour development, leading to an invasive carcinoma, lasts approximately 2 to 5 years as it progresses from mild to moderate to severe dysplasia, and finally to carcinoma in situ (Saccomanno et al., 1974). Squamous cell lung cancer was once the most common type, but there have recently been increases in adenocarcinomas and small cell carcinomas. This may be partly due to improved diagnostic procedures and partly due to risk factors changing over time. All the major types have been associated with smoking, but the association is weaker for adenocarcinomas (Barbone et al., 1997). Non-smokers appear to predominantly develop adenocarcinomas (Wynder and Covey, 1987).

Temporal trends for lung cancer survival rates have remained fairly steady since the early seventies, though a small observed reduction in survival is thought to be due to changing histological patterns (Ugnat et al., 2005). On the other hand, more small cell and large cell cancers have appeared over the last several decades, with comparably fewer squamous cell and adenocarcinoma-type cancers. Evidence suggests differing survival rates, with small cell cancers having the worst five-year survival, at about 5%.

1.2 Etiology of lung cancer

1.2.1 Cigarette smoking

Though tobacco has been used for centuries, the sharp rise in lung cancer rates throughout the twentieth century is thought to be due to the introduction of bulk manufacturing of cigarettes through the early 1900s, the lack of appreciation of the addictive properties of nicotine, and the introduction of additives that increase nicotine availability (DHHS, 1989). These reasons likely ushered in a new pattern of sustained exposure of the lung to inhaled carcinogens. The observation that carcinomas of the lung might be caused by tobacco can be traced back to at least 1898 and other publications throughout the first half of the twentieth century (Doll, 1994). But 1950 brought several sound epidemiologic studies that were the first to clearly demonstrate that tobacco smoke was a strong etiologic factor of bronchogenic carcinomas (Doll and Hill, 1950; Levin et al., 1950; Wynder and Graham, 1950). This relation was later corroborated in studies of several large cohorts (Hammond, 1966; Doll et al., 1994; McLaughlin et al., 1995), but in 1964 the evidence was deemed strong enough for the publication of the landmark Surgeon General's Report that concluded that cigarette smoking caused lung cancer (PHS, 1964).

These studies showed that smokers of two packs or more a day had about a twentyfold higher rate of lung cancer incidence compared to non-smokers. Risk of lung cancer is also strongly dependent on the duration of smoking (Doll and Peto, 1978; IARC, 2004b), with some evidence indicating that the number of cigarettes smoked has less of an effect than the duration of smoking. The risk of lung cancer has been found to decrease after quitting smoking, with the reduction becoming apparent approximately five years after quitting (DHHS, 1990). Regardless of the duration of abstinence, consistently higher risk in ex-smokers has been observed compared to never-smokers (Vineis et al., 2005). A study using a Saskatchewan registry found that women developed cancer at an earlier age than men, even having smoked fewer cigarettes and for a shorter duration (McDuffie et al., 1991). There were also many cases among lifetime non-smokers, 15% and 3% respectively, for women and men.

Findings from several studies suggest that up to 90% of lung cancer deaths are attributable to smoking in the United States (Shopland et al., 1991) and other developed countries (Peto et al., 2005).

Passive inhalation of second-hand smoke has also been related to the incidence of lung cancer (IARC, 2004b). Meta-analyses of studies of non-smoking spouses of smokers provide evidence for a small effect on lung cancer risk, with rate ratios between 1.2 and 1.4 (Pershagen, 1994), with a recent study of environmental tobacco smoke reporting a point estimate and 95% confidence intervals of 1.3 (0.9, 2.0) (Vineis et al., 2005). To approximate passive smoking, it has been estimated that exposure to the equivalent of 0.5 cigarettes a day from birth to 65 years of age, results in a rate ratio of approximately 1.4 compared to lifetime non-smokers (Darby and Pike, 1988).

Smoking is still an active area of research, partly to quantify the complex characteristics of its relation to lung cancer incidence (Leffondré et al., 2002), and partly to study the effects of changes to the cigarette design over the past decades, including reduced tar and nicotine levels (Kabat, 2003).

1.2.2 Other non-occupational causes

Lung cancer is a complex disease that has had several molecular, chromosomal, and cellular events related to its initiation and promotion (Economou et al., 1994). Most of the genetic defects are thought to occur in adulthood, likely due to environmental exposures to carcinogens, but evidence also supports an inherited susceptibility (Shields, 1999). Several clinical reports exist of familial lung cancer (Mulvihill, 1976; Shaw et al., 1991), with results showing about a twofold increase in risk among relatives of someone with lung cancer (Ooi et al., 1986).

Several studies have focused on the theory that lung cancer can develop from the fibrosis or inflammatory fibrotic reaction of previous lung disease, such as tuberculosis (Richardson et al., 1987) or pneumonia (Alavanja et al., 1992; Brownson and Alavanja, 2000). This issue was also raised with respect to asbestos and silica exposures, for which limited evidence exists that asbestosis (Browne, 1986) and silicosis (Ng, 1994), respectively, are on the causal pathway between exposure and lung cancer.

Diet was first linked to lung cancer in a study by Bjelke (1975), in which he reported that lung cancer risk was inversely related to vitamin A consumption. This was later hypothesized to be mainly attributed to beta-carotene (Byers, 1994). While some studies have failed to find an association between diet and lung cancer, the balance of evidence suggests some role for diet. Alberg and Samet (2003) reported that low consumption of carotene, the bottom quartile of use compared to the top quartile, resulted in a 50-100% increase in lung cancer risk. The effect appears to be specific to squamous cell and small cell tumours. Recent work, however, suggests that the supposed effect of beta-carotene observed in non-experimental epidemiologic studies may be due to residual confounding (Stram et al., 2002).

Although the link between consumption of alcohol and lung cancer has often been viewed sceptically because of perceived residual confounding due to smoking history, many recent studies (Bandera et al., 2001) and a meta-analysis (Freudenheim et al., 2005) are suggestive of a small association to lung cancer. In particular, the meta-analysis found the association of alcohol with lung cancer to be approximately 1.2 (0.9, 1.6) in men and even higher in men who had never smoked.

Air pollution is potentially of great concern for risk of lung cancer. While industrial pollutants have been on the decline in recent decades, vehicle engine emissions, including nitrogen oxides and volatile organic compounds, have been on the rise (Boffetta, 2004). A recent review of air pollution studies suggests a rate ratio of about 1.5 for high air pollution urban areas compared to low air pollution rural areas (Boffetta, 2004). Increased risks have been reported in residential areas close to smelters, foundries, and chemical industries. Indoor air has also been implicated, due to fumes released by heated cooking oils (Zhong et al., 1999), asbestos used in indoor building materials (Boffetta and Nyberg, 2003), and build up of radon, a radioactive gas that seeps into homes from soil or groundwater (Lubin and Boice, Jr., 1997).

1.2.3 Occupational etiology

Studies of occupations have given rise to a large body of information on causes of lung cancer and other cancers (Boffetta, 2004). Evidence exists linking many industrial processes and occupational groups and circumstances to a higher risk of lung cancer

(Siemiatycki et al., 2004), but for many substances there has not been adequate evaluation of their carcinogenicity. The present thesis sets out to contribute to such research in the context of an occupational hazard surveillance study conducted in Montreal. The following section reviews the vast literature that exists on many of the currently recognized or suspected occupational lung carcinogens.

2 Literature review: Occupational etiology of lung cancer

This chapter will provide a review of the literature on several currently suspected occupational lung carcinogens. The literature on lung cancer, and even on occupational causes of lung cancer, is extraordinarily large, and so the following reviews are of necessity brief. Following this, typical design and analytic approaches for epidemiologic studies of occupation-related cancer are considered with a view to the methodological improvements made in the context of the present study and thesis.

2.1 Historical roots

Study of environmental agents in the etiology of cancer, whether pollutants at large or substances in the workplace, has only in recent decades come to fruition, providing insight into causal origins and into potential preventive measures. Aside from work on environmental factors, such as ambient air pollution, it was with occupationally exposed populations that most carcinogens were historically first observed (Siemiatycki, 1991). And it is with these typically involuntary and reducible exposures that measures have been taken to improve industrial hygiene, with a view toward controlling cancer (Boffetta, 2004).

Research into the etiology of cancer began in earnest in the twentieth century with animal experimentation and laboratory sciences (Berenblum, 1967). The epidemiologic approach to cancer research, as we know it, has a fairly short history, arguably beginning in the 1950s with landmark studies of the etiology of lung cancer, particularly of smoking (Doll and Hill, 1950; Mills and Porter, 1950; Levin et al., 1950; Schrek and Baker, 1950). Though there were earlier studies of lung cancer and suspicious occupational circumstances, such as nickel refineries (Bridge, 1933), chromate manufacturing (Machle and Gregorius, 1948), and sheep-dip manufacturing (Hill and Faning, 1948), the modern epidemiologic approach to cancer research came to fruition much later and in spectacular fashion. Notable is the demonstration of how large a role that cigarette smoking played in lung cancer etiology and the somewhat smaller but still important role that industrial carcinogens have played (Doll, 1994).

Sir Richard Doll played a prominent role in promoting the perception that lung cancer has an occupational etiology, independent of the role of smoking (Doll, 1959). He published the first quantitative assessment of asbestos exposure and risk of lung cancer (Doll, 1955). Though carcinogens are now recognized as arising from many sources, prior to the 1970s lung carcinogens were substances or circumstances found primarily in the workplace (Siemiatycki et al., 2004).

The International Agency for Research on Cancer (IARC) Monograph Program was initiated in 1969 to evaluate the carcinogenic risk of substances to humans, and to produce comprehensive reviews on individual substances (IARC, 2004a). The program has considered simple chemicals and complex mixtures, as well as other substances, such as radiation and viruses. The program assembles a team of international experts who evaluate the substance from available evidence covering chemistry, animal and laboratory sciences, and epidemiologic research. These experts produce qualitative and quantitative assessments, culminating in a classification of the substance as definitely carcinogenic to humans, probably carcinogenic, possibly carcinogenic, unclassifiable due to inadequate evidence, or probably not carcinogenic. Since its initiation, the IARC monograph program has identified 28 occupational substances as definite carcinogens, 27 as probable carcinogens, and 113 as possible carcinogens (Siemiatycki et al., 2004).

2.2 Known or suspected lung carcinogens

Many occupational circumstances and substances have come under suspicion as being indicators for higher risk of lung cancer. Table 2-1 was abstracted from Siemiatycki et al. (2004), table 7, and Rousseau et al. (2005), table 1. It provides a listing of all the occupational substances, occupations, and industries that are potentially related to lung cancer, and frames the strength of evidence according to the authors' views of strong or suggestive support.

There is a voluminous literature on the occupational etiology of lung cancer, covering hundreds of chemicals and occupational circumstances, hundreds of studies, and dozens of reviews. For the purpose of this literature review, the sections on each chemical are brief, and reviews are provided only for a selection of the occupational substances listed in Table 2-1. Individual studies are not reviewed in the text, but details for many studies

appear in the summary tables for each chemical. These tables are not by any means comprehensive; they are a selection of large and well-known studies, as well as studies involving a diverse sampling of occupations with known exposure to the particular chemical. The layout of the chemical-specific summary tables was drawn, in particular, from a review article by Steenland et al. (1996).

Table 2-1: Occupational substances and circumstances marked as high risk for lung cancer

Strength of evidence	*Substance or circumstance*
Strong	Aluminium production; arsenic and arsenic compounds; asbestos; beryllium; cadmium and cadmium compounds; chromium compounds, hexavalent; coal gasification; coke production; hematite mining, underground, with radon exposure; involuntary (passive) smoking; ionizing radiation; iron and steel founding; selected nickel compounds, including combinations of nickel oxides and sulfides in the nickel refining industry; painters; silica, crystalline; soots; talc containing asbestiform fibres. [for small cell lung cancer only: bis(chloromethyl)ether and chloromethyl methyl ether (technical grade)]
Suggestive	Benz(a)anthracene; benzo(a)pyrene; α-chlorinated toluenes; coal tars and pitches; cobalt metal with tungsten carbide; di-benz[a,h]anthracene; diesel engine exhaust; epichlorohydrin; hairdressers and barbers; inorganic acid mists containing sulphuric acid; inorganic lead compounds; isopropanol manufacture, strong acid process; mineral oils, untreated and mildly treated; non-arsenical insecticides; mustard gas; production of art glass, glass containers, and pressed ware; rubber industry; 2,3,7,8 tetrachlorodibenzo-para-dioxin (TCDD).

2.2.1 Asbestos

Asbestos is a mineral fibre of impure magnesium silicate, often used in fireproofing, insulation, brake linings, and building materials. The two broad classes of asbestos, serpentine (including chrysotile) and amphibole (including amosite and tremolite), are defined by their differing physical structure, where the serpentine structure is curly and the amphibole is needle-like. The differences in structure have lead to the different industrial applications. Animal research has shown that asbestos fibres can cause chromosomal damage in mammalian cells (Walker et al., 1992), and the size of the fibre

is believed to play an important role in its carcinogenicity. Fibres greater than 5 microns in length are thought to be the most dangerous species (Stanton et al., 1981; Wagner et al., 1988).

Table 2-2 provides the results from several published studies of asbestos exposure. According to the respective authors, several of the studies had large enough point estimates to rule out an effect entirely due to confounding by smoking history. Liddel and McDonald (1980) did not adjust for smoking, but did comment on the comparability of smoking distributions in the comparison groups.

Table 2-2: Results from selected studies of asbestos exposure

Reference	*SMR or OR (95% confidence limits)*	*Exposed population*	*Smoking adjustment*
(Selikoff et al., 1979)	SMR 4.1 (3.9, 5.0)	17800 insulation workers	Yes
(Liddell and McDonald, 1980)	SMR 3.5 (2.7, 4.4)	4559 chrysotile miners with asbestosis	Limited
(Newhouse et al., 1985)	SMR 3.0 (2.6, 3.3)	5100 textile plant workers	No
(Seidman et al., 1986)	SMR 5.0 (4.0, 6.0)	820 amosite factory workers	No
(Vineis et al., 1988)	OR 1.2 (0.9, 1.7)	98 cases/90 controls automobile brake workers	Yes
(Newhouse and Sullivan, 1989)	SMR 1.0 (0.9, 1.2)	12571 workers with chrysotile friction products	No
(Neuberger and Kundi, 1990)	SMR 1.0 (0.8, 1.4)	2816 chrysotile cement workers	Yes
(Hughes and Weill, 1991)	SMR 4.3 (2.0, 8.2)	77 exposed men with asbestosis	Yes
(Sluis-Cremer et al., 1992)	SMR 1.7 (1.3, 2.2)	7317 amphibole miners	No
(Hrubec et al., 1992)	SMR 1.1 (0.8, 1.4)	2327 automobile mechanics and repairmen	Yes
(Pira et al., 2005)	SMR 2.8 (2.2, 3.5)	889 men/1077 women asbestos textile workers	No

Asbestos was first quantitatively related to lung cancer among British textile workers in a study by Doll (1955). Epidemiologic evidence has accumulated over the past decades, showing that the risk of lung cancer is increased in several asbestos-related industries, including miners, textile workers, friction product workers, insulation workers, shipyard workers, and cement workers (McDonald and McDonald, 1987). While some agencies have claimed all types of asbestos are carcinogenic and that there should be a worldwide ban on its use (LaDou et al., 2001), and compelling evidence exists for all forms of asbestos being carcinogenic for one animal species or another (Doll and Peto, 1985), there is still disagreement as to whether present-day exposure to both classes are equally culpable as causes of human lung cancer (Camus, 2001; Siemiatycki, 2001).

IARC classified asbestos as a definite human carcinogen in 1977, backed by sufficient evidence from both animal and human research (IARC, 1977). Individuals with asbestosis (implying high exposure to asbestos) appear to be more prone to lung cancer. A meta-analysis of six studies of asbestotics produced a summary relative risk and 95% CL of 5.9 (5.0, 7.0) and an analysis of twenty cohort studies of asbestos workers produced a summary relative risk estimate of 2.0 (1.9, 2.1) (Steenland et al., 1996). Disentangling the relationships of asbestos exposure, asbestosis, smoking, and lung cancer has been difficult (Hessel et al., 2005). Some studies have suggested that fibrosis, such as from pre-existing asbestosis, is a necessary antecedent to lung cancer (Liddell and McDonald, 1980; Hughes and Weill, 1991). If this is indeed the case, and there was a threshold whereby asbestosis did not occur at the lower levels of exposure common among current workers (Weill, 1994), then there would be implications as to whether current (low) asbestos exposure should still be considered hazardous. That is, if asbestosis is a necessary intermediate pathology, it would imply that typical current levels of asbestos exposure are unlikely to cause lung cancer. Recent opinion holds that epidemiologic evidence alone is not sufficient to settle the issue (Stayner et al., 1997).

A decline of risk has been associated with increasing time since last exposure to asbestos, suggesting that asbestos may act as a late-stage promoter (Seidman and Selikoff, 1990). Finally, though evidence has supported the combined dependence of the effects of asbestos exposure and cigarette-smoking, there have been conflicting claims as to whether

these joint effects are closer to multiplicative or additive (Berry et al., 1985; Liddell and Armstrong, 2002).

2.2.2 Silica

Silica is a crystalline compound commonly found in the forms of quartz and sand. Exposure to crystalline silica is common in miners, depending on ore content, and in industries of masonry, stonework, pottery, and glasswork. Animal studies offer conflicting evidence of its lung carcinogenicity across different rodent species (Saffioti, 1992). Nevertheless, there was enough evidence for IARC to classify silica as a probable carcinogen in 1987 (IARC, 1987b; IARC, 1997), based on what was termed sufficient evidence in animals and limited evidence in humans, with the latter due to often inconsistent results from studies.

Table 2-3 provides the results from several published studies of either silica exposure or patients with silicosis. Of note, in the study reported by Steenland and Brown (1995), the effect of silica was likely unconfounded by exposure to arsenic and radon in miners.

Studies of lung cancer in relation to silicosis, a respiratory illness resulting from high exposure to silica, have provided evidence of a stronger relation than have studies of actual silica exposure. This proves somewhat difficult to interpret, as it may be that silicosis itself induces lung cancer. Exposure to silica is often accompanied by exposure to PAHs, radon, asbestos, and several other suspected carcinogens (Blot and Fraumeni, Jr., 1996). Steenland et al. (1996) provide a best evidence synthesis and meta-analysis of fifteen studies of silicotics [summary relative risk and 95% CL of 2.8 (2.5, 3.2)] and thirteen studies of silica-exposed workers [summary relative risk of 1.3 (1.2, 1.4)], both of which excluded studies of mines with known confounding exposures, studies using autopsy-based information, and proportionate mortality studies with possible selection biases.

The judgement by IARC has remained somewhat controversial as several studies have not demonstrated an increase in risk. This might be explained by recent work, which suggests that the exposure-response curve of silica and lung cancer is relatively low, implying that a substantial number of highly exposed individuals would be necessary to detect the effect (Steenland et al., 2001).

Table 2-3: Results from selected studies of silica exposure

Reference	*SMR (95% confidence limits)*	*Exposed population*	*Smoking adjustment*
(Costello and Graham, 1988)	SMR 1.2 (1.0, 1.4)	5414 granite workers employed '50 to '82	No
(Infante-Rivard et al., 1989)	SMR 3.5 (3.1, 3.9)	1165 compensated silicotics	No
(Koskela et al., 1990)	SMR 1.6 (1.0, 2.2)	1026 granite workers	No
(Ng et al., 1990)	SMR 2.0 (1.4, 2.9)	1419 men in silicosis registry	Yes
(Winter et al., 1990)	SMR 1.3 (1.0, 1.7)	3669 pottery workers	Yes
(Merlo et al., 1991)	SMR 1.5 (1.0, 2.1)	1022 brick workers	Yes
(Amandus and Costello, 1991)	SMR 1.3 (0.8, 2.0)	724 silicotics	Yes
(Steenland and Brown, 1995)	SMR 1.1 (0.9, 1.4)	3328 gold miners	Yes
(Checkoway et al., 1997)	SMR 1.3 (1.0, 1.6)	2342 diatomaceous earth miners	No

2.2.3 Diesel engine exhaust

Diesel engine exhaust is a complex and variable mixture of chemicals, composed of several thousand oxidation and nitration products resulting from the incomplete combustion of diesel fuel (IARC, 1989). The relative proportions of the mixture depend on the type of engine, the operational circumstances, the fuel composition, and the presence of any emission controls. Many of the chemicals overlap with the constituents of tobacco smoke. Exposure is common among truck drivers and construction workers, but moderate levels can also be found in the general environment. It is the particulate phase that has been implicated in carcinogenicity (IARC, 1989). Diesel exhaust has gradually been restricted in North America with tighter vehicle emission standards, but exposure limits for nitrogen dioxide, sulphur dioxide, carbon monoxide, and the particulate phase, differ from one Canadian province to another.

Many of the constituent chemicals of diesel have been assessed by IARC and deemed possibly carcinogenic for some cancers: acetaldehyde (possible carcinogen), benzene (definite carcinogen), formaldehyde (probable carcinogen), lead compounds (possible

carcinogen), and benzo(a)pyrene (probable carcinogen). In 1989, IARC classified whole diesel engine exhaust as a probable carcinogen (IARC, 1989).

Table 2-4 presents the results from some major studies of diesel engine emissions. There are few studies with evidence of clear relevance for evaluating the causal role of whole diesel exhaust in lung cancer, and many of the existing studies are hampered by the lack of data on relevant exposure. Focussing on six studies, a meta-analysis by Steenland et al. (1996) provided a combined relative risk estimate and 95% CL of 1.3 (1.1, 1.4). Other engine emissions, such as gasoline engine exhaust, have not been properly assessed due to insufficient evidence (IARC, 1989).

Table 2-4: Results from selected studies of diesel exhaust exposure

Reference	*SMR, OR, or HR* [a] *(95% confidence limits)*	*Exposed population*	*Smoking adjustment*
(Boffetta et al., 1988)	SMR 1.2 (1.0, 1.4)	American Cancer Society Cohort; 62800 self reported diesel exposure	Yes
(Gustavsson et al., 1990)	SMR 1.2 (0.7, 1.8)	695 bus garage workers	No
(Steenland et al., 1990)	OR 1.9 (1.0, 3.4)	56 cases/36 controls Long-haul diesel truck drivers	Yes
(Boffetta et al., 1990)	OR 1.2 (0.7, 2.0)	35 cases/49 controls Workers exposed to diesel exhaust	Yes
(Emmelin et al., 1993)	OR 1.7 (1.4, 2.1)	50 cases/154 controls Shipyard workers	Yes
(Garshick et al., 2004)	HR 1.3 (1.1, 1.6)	204 cases Long term railroad workers	No

[a] HR, hazard ratio

2.2.4 Beryllium

Beryllium is a metallic element used in aerospace materials, in nuclear reactors, and in copper alloys for springs and electrical contacts. Exposure to beryllium occurs mostly in mining and in the manufacturing of ceramics and electronic equipment. Lung cancer has been induced in rats and monkeys by beryllium exposure (Groth, 1980), but there have been few epidemiologic studies (see Table 2-5 for a selection of studies).

In 1993, IARC classified beryllium as a definite carcinogen (IARC, 1993), largely based on the results of two studies: Steenland and Ward (1991) and Ward et al. (1992). One of the seven processing plants in the latter study was followed further and the data re-analyzed in Sanderson et al. (2001).

Table 2-5: Results from selected studies of beryllium exposure

Reference	*SMR (95% confidence limits)*	*Exposed population*	*Smoking adjustment*
(Mancuso, 1980)	1.4 (1.1, 1.7)	3685 beryllium plant workers	No
(Infante et al., 1980)	1.9 (0.8, 4.0)	421 men in beryllium registry	No
(Hinds et al., 1985)	1.7 (0.8, 3.5)	19 cases/17 controls Self-reported exposure	Yes
(Steenland and Ward, 1991)	2.0 (1.3, 2.9)	689 patients with berylliosis	Limited
(Sanderson et al., 2001)	1.2 (1.0, 1.4)	3569 beryllium plant workers	Limited

2.2.5 Chromium

Chromium is a metal commonly found in trivalent and hexavalent oxidative states. Its various compounds had early usage in the manufacturing of inorganic pigments, but its most important use has been as an alloy for chrome plating and stainless steel. There have been several epidemiologic studies since the first observation among workers in U.S. chromate-production factories (Machle and Gregorius, 1948). IARC classified insoluble chromium VI as a definite carcinogen in 1990, based on what it deemed sufficient

evidence in animals and in humans (IARC, 1990). Evidence of human carcinogenicity has been inconsistent for soluble chromium III and controversy remains as to the relative effects of the trivalent/hexavalent forms (Mancuso, 1997a; Mancuso, 1997b; Gibb et al., 2000).

Table 2-6 provides the results from several published studies of chromium exposures. In a meta-analysis of ten large and well-designed studies of chromium workers, manufacturers of chromate paints, and chromate plating workers, the combined summary estimate and 95% CL for lung cancer was 2.8 (2.5, 3.5) (Steenland et al., 1996). Many of the existing epidemiologic studies, however, documented historical chromium exposures that were tenfold higher than modern levels (IARC, 1990). Limited evidence of the effects of more recent chromium exposures suggests that the risk of lung cancer has decreased, possibly due to improvements in occupational hygiene from modifications of industrial processes (Luippold et al., 2005).

Table 2-6: Results from selected studies of chromium exposure

Reference	*SMR (95% confidence limits)*	*Exposed population*	*Smoking adjustment*
(Enterline, 1974)	9.4 (7.3, 11.9)	1212 chromate plant workers	No
(Hayes et al., 1979)	2.0 (1.6, 2.6)	1850 chromium plant workers	No
(Frentzel-Beyme, 1983)	2.0 (1.2, 3.2)	Workers in a chromate pigment plant	No
(Davies, 1984)	1.8 (1.4, 2.4)	1152 chromate pigment workers	No
(Sorahan et al., 1987)	1.5 (1.2, 1.9)	2689 chromium/nickel platers	No
(Hayes et al., 1989)	1.4 (0.9, 2.1)	1879 chromium pigment workers	No
(Takahashi and Okubo, 1990)	1.9 (0.8, 3.7)	626 chromium platers	Limited
(Korallus et al., 1993)	1.3 (0.6, 2.4)	678 chromate production workers (post modification)	No
(Luippold et al., 2005)	0.8 (0.2, 2.4)	617 chromate production workers (post modification)	No

2.2.6 Nickel

Nickel and its salts are used in electroplating and in the production of stainless steel, nonferrous alloys, and batteries. Studies of nickel alloy manufacturers provided inconsistent evidence of increased lung cancer risk (ICNCM, 1990), and many of the published results are likely confounded by other exposures, such as chromium VI, arsenic, and PAHs (Blot and Fraumeni, Jr., 1996). An early study by Doll (1977) found that the increased risk in refineries was limited to the earliest years of operation, when exposure was much heavier than current (even by 1930) standards. However, using thirteen studies of workers in nickel refineries, a recent meta-analysis provided a summary relative risk estimate and 95% CL of 1.6 (1.4, 1.7) (Steenland et al., 1996).

IARC classified nickel compounds (in general) as definitely carcinogenic to humans based on animal and human evidence, but metallic nickel in particular was only classified as possibly carcinogenic (IARC, 1990). Recent reviews and the report by the International Committee on Nickel Carcinogenesis in Man (1990) came to the same conclusion that the strongest evidence supports a carcinogenic role for soluble nickel and nickel oxides and sulphides, but not metallic nickel in non-refining processes (Hayes, 1997).

Table 2-7: Results from selected cohorts reviewed in ICNCM (1990)

Reference	*SMR (95% confidence limits)*	*Exposed population*
INCO (Canada)	3.1 (2.4, 4.0)	1754 nickel refinery workers
Falconbridge (Canada)	1.3 (1.0, 1.6)	11595 nickel refinery workers
Hanna Nickel Smelting Company (United States)	1.4 (0.9, 2.3)	1510 nickel miners and smelter workers
Henry Wiggin Alloy Company (England)	1.0 (0.6, 1.5)	1907 nickel alloy manufacturers

2.2.7 Arsenic

Arsenic is a metallic element that exists in both inorganic and organic forms, and has been used in insecticides, weed killers, and various alloys. Miners, workers at copper smelters, manufacturers of sheep dip compounds, and those manufacturing certain pesticides have been substantially exposed to arsenic in the past, and are thought to be at higher risk for lung cancer (Hayes, 1997). Though animal studies have not provided adequate support for arsenic's role as a carcinogen, the evidence in humans has been consistent (Blot and Fraumeni, Jr., 1994). IARC classified arsenic as a definite carcinogen (IARC, 1980).

Table 2-8 presents results from selected studies of arsenic exposure. In a meta-analysis of six studies, Steenland (1996) derived a summary relative risk estimate and 95% CL of 3.7 (3.1, 4.5). Some evidence suggests that arsenic may act as a late stage promoter of lung tumours (Brown and Chu, 1983). And the few studies that also assessed smoking habits indicate that cigarette smoking and arsenic may act synergistically (Pershagen et al., 1981; Jarup and Pershagen, 1991).

Table 2-8: Results from selected studies of arsenic exposure

Reference	*SMR or OR (95% confidence limits)*	*Exposed population*	*Smoking adjustment*
(Ott et al., 1974)	SMR 3.4 (2.1, 5.3)	603 men in pesticide plant	No
(Lee-Feldstein, 1986)	SMR 2.8 (2.6, 3.2)	8045 copper smelter workers	No
(Enterline et al., 1987b)	SMR 1.3 (1.1, 2.6)	6078 copper smelter workers	Yes
(Taylor et al., 1989)	OR 15.2 (4.9, 52.7)	Tin miners	Yes
(Jarup et al., 1989)	SMR 3.7 (3.0, 4.5)	3916 copper smelter workers	Yes, in (Jarup and Pershagen, 1991)
(Lubin et al., 2000)	SMR 1.6 (1.2, 2.0)	8014 copper smelter workers	No
(Binks et al., 2005)	SMR 1.6 (1.4, 1.7)	1462 tin smelter workers	No

2.2.8 Cadmium

Cadmium is a metallic element commonly used in electroplating, alloys, solders, dental amalgams, battery contacts, and occasionally fertilizers. Although cadmium usage in the European Union has declined, world wide usage has risen (Jarup, 2003). Lung tumours have been induced in rats by cadmium exposure (Heinrich, 1992), but evidence in humans has been hard to come by due to the common co-occurrence of other exposures in the workplace, including nickel and arsenic (Boffetta, 1992). The low-level exposure of cadmium from cigarette smoking has been estimated to account for 1 to 18 lung cancer deaths per 10000 smokers (Hertz-Picciotto and Hu, 1994). IARC classified cadmium exposure in 1993 as a definite carcinogen in humans (IARC, 1993).

Table 2-9: Results from selected studies of cadmium exposure

Reference	*SMR (95% confidence limits)*	*Exposed population*	*Smoking adjustment*
(Kazantzis et al., 1988)	1.2 (1.0, 1.3)	6995 Cadmium processing plant workers	No
(Stayner et al., 1992)	2.7 (1.2, 5.1)	606 Cadmium smelter workers	No
(Sorahan and Esmen, 2004)	1.1 (0.8, 1.5)	926 Nickel-cadmium battery manufacturers	No

2.2.9 Polycyclic aromatic hydrocarbons

Polycyclic aromatic hydrocarbons (PAH) are a complex group of chemicals composed of variable numbers of benzene rings, and formed through the incomplete combustion of organic material. The profile of PAHs is very much dependent on the conditions of their production, and exposure is widespread in the human environment. They are characteristic of, among other things, engine emissions, tobacco smoke, and cooking fumes. Since humans are never exposed solely to individual PAHs, but rather to complex mixtures, animal evidence has played a crucial role in assessing carcinogenicity (IARC, 1983). Further, distinguishing the effects of the different species of PAHs from the effect of the substances they are adsorbed to, such as soot, is neither easy nor accomplishable in

an obvious fashion. Epidemiologic evidence does not exist to directly address the role of PAHs, so instead studies have assessed the chemicals or occupations in which PAHs are present to a high degree, addressing perhaps the question of what proportion of the effect can be attributable instead to PAHs (IARC, 1983).

Early suspicions about PAHs came from studies of the excess risk of lung cancer found in gas workers exposed to coal carbonization products (Doll et al., 1972). A portion of the increases in risk of lung cancer among coke oven workers can plausibly be accounted for by PAHs since benzo(a)pyrene-DNA adducts have been detected in blood samples of the workers (Harris et al., 1985). Other occupations and industries, for which PAHs have been implicated, include steelworkers, workers in aluminium smelters, and roofers exposed to asphalt and pitch (Blot and Fraumeni, Jr., 1996). The increased risk of lung cancer found in these occupations, as well as in workers exposed to diesel engine exhaust, is thought to be due to PAH exposure and not the other chemicals common to those workplaces (Boffetta et al., 1997).

2.2.10 Radon gas

Radon is a naturally occurring radioactive noble gas. Of the various isotopes, it is the decay from uranium that is a common exposure in the human environment. Following inhalation, the irradiation of respiratory tissue can cause consequent chromosomal damage (Jostes, 1996). While studies of uranium miners have consistently demonstrated an excess risk of lung cancer (BEIR IV, 1988), the low concentrations found in poorly ventilated buildings and homes is a major concern because of the numbers of individuals exposed and the persistence of exposure (Clarke and Southwood, 1989). A recent review by IARC (2001) combined eight case-control studies of residential radon, resulting in a summary RR and 95% CL of 1.1 (1.0, 1.2), for exposures of 100 Bq/m^3. Darby (2001) has estimated that up to 2000 lung cancer deaths a year in the United Kingdom are caused by residential radon exposure. IARC has classified radon as a definite carcinogen in humans (IARC, 1988).

2.2.11 Other substances

What follows is brief survey of several other substances that are currently, or were historically, under consideration as lung carcinogens.

The evidence for the carcinogenic effects of acrylonitrile arose from animal research and some limited evidence in humans. In contradiction to the initial study by O'Berg (1980), more recent evidence (Collins et al., 1989) has not supported a lung carcinogenic effect of acrylonitrile, including follow-ups of the initial cohort (Chen et al., 1988a; Chen et al., 1988b). IARC did classify acrylonitrile as a probable carcinogen based on sufficient animal evidence (on sites other than lung) and limited human evidence (IARC, 1979).

Chloromethyl ethers were first discovered to be carcinogenic in laboratory animals, and human studies demonstrated their role in small cell lung carcinomas (Collingwood et al., 1987; IARC, 1987a), apparently having an effect in both early and late stages of carcinogenesis. Vinyl chloride, common in early refrigerants and aerosol propellants, has been linked to lung cancer in several studies (Buffler et al., 1979; Wu et al., 1989), and classified as a definite carcinogen (IARC, 1987a).

Rubber workers involved in tire curing and fuel cells have been reported as being at higher risk for lung cancer (Monson and Fine, 1978). Studies have found workers at increased risk due to ferric oxide dust and ferrochromium, but not ferrosilicon (Axelson and Sjoberg, 1979; Langard et al., 1980). Welders in shipyards have been implicated (IARC, 1990), as have workers with mild steel dust (Moulin et al., 1993). Elevated risks have also been reported among workers of steel pickling and operations where exposure to sulphuric acid and other acid mists is relatively high (IARC, 1992). Lead exposure has been inconsistently related to lung cancer, though a meta-analysis of several studies suggests a low effect of 1.1 (1.0, 1.2) (Steenland and Boffetta, 2000b). The evidence from several studies of formaldehyde remains inadequate (Blair et al., 1986; IARC, 1995). As much as threefold risk of lung cancer due to toluene has been found in several small studies (Wong, 1988; IARC, 1992).

Several other industrial exposures are suspected of being lung carcinogens, though evidence is still preliminary for most (see recent reviews by Blot and Fraumeni, Jr. (1996), Boffetta (2004), and Siemiatycki et al. (2004)).

2.3 Undiscovered carcinogens

There are a number of chemicals identified in animals as lung carcinogens but that have remained unstudied in humans, such as ceramic fibres, acrylamide, dichloromethane, and

acetaldehyde (Boffetta, 2004). Of the many chemicals in common usage in industries, relatively few have been studied for their health consequences. And a complex series of coincidences were necessary for clinicians of the past to discover many of those carcinogens of which we are currently aware (Siemiatycki, 1991). These factors lead to a reasonable suspicion that there remain several more undiscovered carcinogens, and with a workplace that is constantly evolving, new chemicals are introduced regularly.

2.4 Consideration of methods for studying occupational cancer

2.4.1 Study designs

Most human carcinogens found in the workplace were initially suspected through observation by astute physicians (Doll, 1975; Siemiatycki, 1984). These clinical discoveries were somewhat haphazard, requiring several coincident events: the tumours needed to occur in a large enough group of people with some shared occupation or workplace; the tumours needed to have been identified in a single clinic; and the tumours needed to occur with sufficient incidence to be noticeable. Epidemiologic study design is a step toward removing many of these 'conditions.'

Broadly speaking, there are three main sources of epidemiologic information pertaining to occupational exposures and cancer:

1. Routine record analyses, such as tumour registries and death certificates,
2. Occupational cohort studies, typically with retrospective mortality information,
3. Case-control studies.

There has been extensive use of routine records, such as death certificates, to relate occupations to cancer mortality with either a proportionate mortality ratio or a standardized mortality ratio. While this is an inexpensive approach, it suffers several disadvantages, such as inaccurate job labelling, incorrect diagnostic information, incomplete job history, and lack of information on confounders, apart from sex and age. Generally, mortality is not a good indicator of incidence. While such criticisms are partly alleviated by the use of tumour registries, work history is rarely recorded.

One mainstay of occupational epidemiology has been the opportunistic use of cohorts of workers with a shared exposure to a substance. Such studies have been useful when the

case was made that a particular work cohort had relatively few other 'confounding' workplace exposures. This approach is based on the availability of databases linking work history and health events. For use in epidemiologic research, establishing prospective systems of data collection would require decades of wait due to the nature of carcinogenicity. While existing data systems have been used, they often suffer from incompleteness and a lack of important non-occupational information. Further, exposures in previous employment would typically be unavailable, often necessitating the restriction to studying long term employees.

The typical design of an occupational cohort study compares the rates in the work cohort to national, population rates, using age and sex standardized incidence ratios or mortality ratios. Concern here would be with bias due to the nature of this comparison, which involves selectively healthy workers (Wang and Miettinen, 1982). Internal cohort comparisons can remedy this problem. Rarely is there control for the confounding effects of smoking, which can pose serious problems for studies of lung cancer, though evidence suggests that the resulting bias still allows for the detection of risky substances under certain circumstances (Siemiatycki et al., 1988).

By comparison, the case-control design allows for an often efficient means to ascertain incident cancers and to collect lifetime work history on multiple occupational exposures and potential confounders. The documentation of lifetime work history directly from study subjects has been used effectively to survey whole ranges of occupations and industries (Siemiatycki, 1984). There are several instances of investigators conducting case-control studies within occupational cohorts, but the more typical design has used the catchment population of clinics or hospitals. Identifying incident cases of cancer through medical sources offers greater diagnostic validity than information on death certificates.

Few such studies involve direct hygiene measurements, and the challenge has been to devise improved methods of retrospective exposure assessment (Teschke et al., 2002).

2.4.2 Documenting exposure

Direct industrial hygiene measurements are at face-value the gold standards of exposure assessment, though measures of dose at the target organ would be of even greater scientific relevance (Checkoway et al., 2004). While the direct exposure assessments

involve real time measurement in the workplace, unfortunately, measurements on historical work cohorts rarely exist or are incomplete, and designing a prospective system of hygiene measurement is prohibitively expensive for most studies (Siemiatycki, 1984). However, when direct exposure measurement is possible, the resulting quality of information on the measured exposures can far surpass that which is typically found by other methods.

There are several other approaches to measuring work related exposures. Some have tended to align with particular study designs. For example, occupational cohort studies have traditionally relied on identifying work cohorts with a fairly narrow range of chemical exposures, and exposure to a particular substance is assumed given the employment in that workplace. Identifying high risk occupations has been fruitful, but arguably, characterizing actual exposure histories as opposed to job histories should improve identification of likely carcinogens (Siemiatycki, 1984).

While checklists of chemicals have been used during interviews and self-administered questionnaires to help elicit information about lifetime exposure history directly from the study subject, this method is difficult to implement if the number of chemical exposures is too large (Siemiatycki, 1984). It further suffers from workers not necessarily being familiar with all the naming conventions of the substances in their work environment.

At its simplest level, attributing a chemical exposure to a group of workers with a common job title or within a common industry has been the most common form occupational cancer research in epidemiology. Alternatively, a job-exposure matrix is an algorithm that allows for the translation of occupational or industrial titles into a set list of exposures. This allows for the analysis of exposures irrespective of the occupation they occurred in, and its formation can be via expert opinion (Teschke et al., 2002), observation during walkthroughs (Sieber et al., 1991), or direct hygiene measurements (Kauppinen et al., 1998). Hoar et al. (1980) pioneered the effort to create a job-exposure matrix. Her matrix involved 500 job classes and 376 suspected carcinogens or toxins.

One problem with job exposure matrices is that they do not tend to be transportable to other countries, and they fail to take into account the temporal changes in work place hygiene (Siemiatycki, 1984). Another difficulty with this approach is that, insofar as

chemical carcinogenicity is the concern, occupational labels can be misleading because chemical exposures can be vastly different among workers with similar job titles (Siemiatycki, 1984).

Another method of exposure assessment, developed by Siemiatycki (1981), involves expert-based coding. This method has the work history of each study subject translated into occupational exposures by chemists and hygienists (Siemiatycki, 1984). Based on information elicited from a study subject's interview, from technical documents, and from industry consultants, experts can make informed guesses about probable chemical exposures in the subject's past. As opposed to the need to create a job-exposure matrix prior to its use, the expert-based system is tantamount to the gradual creation of a matrix as the need arises for particular jobs or particular substances, all the while taking account of temporal changes and the subtleties and complexity of work place exposures and tasks. Limited evidence suggests that this approach provides for accurate and reliable exposure measurements (Goldberg et al., 1986; Fritschi et al., 1996; Siemiatycki et al., 1997).

2.4.3 Analytic methods

In assessing cancer and occupational exposures, the extent of handling confounding due to non-occupational characteristics such as age, sex, and smoking history, varies widely across publications. A large part of occupational cancer evidence arose from work cohorts compared to national populations, and these studies have tended to be limited to standardization of age and sex distributions. Few have sufficient information to account for the complexity of the confounding that might be due to smoking, though recent work suggests that even minimal adjustment for current smoking status is satisfactory (Richiardi et al., 2005). With the tendency to have more available information, analytic methods found in case-control studies have provided for greater control of confounding from non-occupational determinants of cancer risk.

On one hand, rarely is mutual confounding considered among occupational chemicals. Given that workers tend to have multiple exposures in their work history, evincing complex correlation patterns in the work population, there is much opportunity for overestimation of effects that are truly null. On the other hand, studies that do document multiple exposures are often not sufficiently large to allow for simultaneous estimation,

which would take into account any mutual confounding. The predominant strategy in studies of multiple exposures has been to address each exposure in its own tailored regression model, using some method for the pre-selection of confounding variables.

The approaches to handling multiple exposures can be summarized by the following three strategies (Greenland, 1993):

1. A single, so-called full model with all the study's variables. Such models have been preferred over any sort of pre-selection strategy, though they are likely to be difficult to fit because of sparse data (Kleinbaum et al., 1982; Miettinen, 1985; Rothman and Greenland, 1998); and if fit, can result in biased and imprecise estimates (Witte and Greenland, 1996).
2. A single model with a reduction of parameters by some pre-selection. This approach implicitly assumes that the variables deleted from the model do not represent true determinants of risk of disease (Greenland, 1989a). Several authors have expressed concern with such models because the final estimation depends on the results of the pre-selection (Sclove et al., 1972; Greenland, 1989a), often leading to inappropriately narrow confidence interval estimates and overstated strength of the evidence (Viallefont et al., 2001).
3. A separate regression model for each parameter of interest. Each model is tailored to one parameter, and only confounders are added or retained in the model, producing an ad hoc compromise between strategies one and two.

Strategies for the pre-selection of variables to include in a regression model, whether using statistical significance or one of several other options, stem largely from the tradition of modeling having solely to do with the addition or removal of confounders (Mickey and Greenland, 1989; Maldonado and Greenland, 1993), which some think is not an appropriate guide for situations with multiple determinants of interest (Greenland, 1993). Another criticism of pre-testing is that sequential testing of single parameters fails to appreciate that confounding can occur by the aggregate of many "small confounders" (Miettinen, 1976b), statistically significant or not, and that these cannot be identified one at a time with any reasonable certainty.

Though the full model strategy to address confounding is not a realistically viable option because of the expected instability of the estimates, Witte et al. (1994) use this model as a step toward "more stable" estimates in the context of multilevel models (Greenland, 2000a). The single model arguably has the benefit of simplicity and parsimony, when compared to K separate models and the procedures that went into their production. On the other hand, it may be unattractive because of the assumptions necessary to believe a large-parametered model, where the data would never have been sufficient to stratify on so many variables. The suggested formal compromise is the addition of a second-level model to the first-level "full model" (Greenland, 2000a). This approach is otherwise known as empirical Bayes modeling, and it can improve estimation of the effects of the entire set of exposures under study and handle large models that ordinarily would not be efficiently estimated.

3 Literature review: Empirical Bayes models

The chapter outlines a Bayesian approach to analyzing multiple exposures, typical of complex exposure histories. It will begin with the multiple inference problem, which will serve as the motivation for introducing a family of analytical techniques known as empirical Bayes models, ultimately leading to the semi-Bayes variant.

3.1 The multiple inference problem

Epidemiologic studies often harvest information across a wide range of human experience, including diet, environment, behaviour, genetics, and, as in the present context, occupational circumstances. The statistical analysis of such datasets engenders problems of inference and method that can be referred to as multiple comparisons or multiple inferences. Perform enough comparisons and unusual results will eventually be found (Tukey, 1991). Savitz and Olshan (1995) claim that it is irrational to criticize a study because of the number of exposures assessed. Some authors have denied that the problem even exists (Rothman, 1990), while others have sought to deal with these issues in a fashion more suited to epidemiologic research than the traditional methods (Greenland and Robins, 1991).

One classical approach to multiple inference arose in the context of significance testing and analysis of variance, where in a study of K groups, there would be a possible K(K-1)/2 pair-wise comparisons (Hsu, 1996). While the probability of rejecting any given null hypothesis should not depend on other comparisons being conducted, the probability of falsely rejecting at least one true null hypothesis nevertheless increases with each additional comparison. In fact, if all the K hypotheses are mutually independent and all K null hypotheses are true, then the overall probability of falsely rejecting at least one hypothesis at α significance level is $1-(1-\alpha)^K$. The desire to maintain a fixed overall error rate α for the statistical significance tests of an entire study involving K comparisons, led to the development of the Bonferroni correction: if there are K independent tests, the corrected significance level is α/K. Epidemiologists have criticized such an adjustment of the significance level because, on the one hand, it can be overly stringent in studies that involve a multitude of exposures, and, on the other hand, significance levels may not be

the appropriate focus for multiple inference problems in epidemiologic studies (Savitz and Olshan, 1995; Greenland, 2000b).

Inference problems can be distinguished by whether they are motivated by single inference questions (is this determinant related to this illness?) or by multiple inference questions (are any determinants in this set related to this illness?) (Rothman and Greenland, 1998). The latter question evokes the universal null hypothesis (no determinants in this set are related to the illness) and an omnibus alternative hypothesis (at least one determinant in this set is related to the illness), which has been criticized as a poor research strategy for most epidemiologic problems (Rothman, 1990; Savitz and Olshan, 1995). There has been a range of opinions expressed on the need for adjustment for multiple comparisons in different contexts (Poole, 1991; Wacholder et al., 2004). However, when the objectives of a study include decision making based on the results, such as in prioritizing further research, then recent opinion suggests that there exist different shades of the multiple inference question and that this can be addressed by non-Frequentist methods of analysis (Greenland and Robins, 1991). Individual hypotheses are maintained but with the qualification that they belong to a set (which of this list of determinants is related to this illness?). Such formulations have been suggested as appropriate for data searches or surveys of multiple determinants (Rothman and Greenland, 1998), such as the occupational cancer study that formed the basis of the present thesis (Thomas et al., 1986). This outlook has led to the development of estimators that on average improve estimation over maximum likelihood estimates for a set of determinants (Efron and Morris, 1973).

Some epidemiology theoreticians have rejected the view that the evidence for a particular hypothesis is affected by how many other hypotheses are considered in the study (Miettinen, 1985; Rothman and Greenland, 1998). One aspect of their criticism is that it makes little sense to require different inferences from two investigators with the same data simply because one considered some additional comparisons within that data (Thomas and Clayton, 2004). The countervailing suggestion is that it is the prior credibility of an hypothesis that determines how results are interpreted (Cole, 1979; Miettinen, 1985; Rothman and Greenland, 1998): The lower the prior credibility, the stronger the evidence that is required. This naturally leads to a Bayesian framework, and

it is of relevance to current genetic epidemiology studies with millions of comparisons being feasible in a given database (Thomas and Clayton, 2004). As opposed to the concern for error frequencies common in the Bonferroni method of multiple inference, empirical Bayes models have been developed as an alternative for *estimation* problems that arise in multiple inference settings (Greenland, 2000b).

The introduction of empirical Bayes models, which are discussed below, is based on three distinct lines of reasoning. The first is that while a single maximum likelihood estimate is asymptotically unbiased, maximum likelihood estimates can be paradoxically biased when considered in an ensemble (Thomas, 1985); that is, when considering a set of point estimates, the largest estimate in the set is more likely to be an over-estimate than an under-estimate of its true value. And imprecise estimates are more likely than precise estimates to be outliers. Both of these phenomena invoke the concept of regression to the mean, which naturally leads to a desire to shrink estimates closer to the center of the distribution of maximum likelihood estimates, with the amount of shrinkage possibly in proportion to their imprecision (Thomas, 1985; Steenland et al., 2000). The second reason is that there is a known bias in the maximum likelihood estimator that occurs in sparse data (Cordeiro and McCullagh, 1991), which, according to Greenland (2000c), occurs more frequently than is commonly thought. This second reason also provides a rationale for the shrinkage estimator because the bias is predictably away from the null value (Cordeiro and McCullagh, 1991). The third reason is that there has been increasing attentiveness to the advantages of Bayesian analyses in epidemiology, with the concept of prior credibility as the basis for evaluating evidence (Miettinen, 1985; Rothman and Greenland, 1998). Within a study of multiple exposures, however, our knowledge may not be rich enough to set a complete Bayesian prior on each parameter (Robert, 2001). Recognizing that particular sets of similar exposures might have similar effects, leads to the concepts of correlated hypotheses and exchangeable parameters.

3.2 Exchangeability and empirical Bayes models

To illustrate empirical Bayes estimation, I will use the hypothetical example of a study in which data have been collected on five types of chromate compounds (ammonium chromate, barium chromate, basic lead chromate, calcium chromate, and potassium

chromate), with the objective of estimating each chemical's effect on risk of illness, represented by the log odds ratio, B_i. This explanation derives from similar examples given by Greenland (1999; 2000a).

When several parameters are considered for estimation, such as for the individual effects of the five chromate compounds, the common approach would be to estimate the value of each parameter, B_i, from the data using a maximum likelihood estimator. Each parameter can only be estimated if there is sufficient data for the estimation. Some of the estimates may be precise and others so imprecise as to render the estimates useless. Given that these are similar chemicals, if the data is not sufficient to estimate each of the parameters separately, a more stable estimate might be attained by pooling the data and assuming each chemical's effect is equivalent, $B_1 = B_2 = B_3 = B_4 = B_5 = \overline{B}$, in this case represented by a single parameter for the chemical category, "chromate compounds." An *ad hoc* approach might assess risk in relation to the whole category first, and if appropriate, follow this with assessments of the individual elements. This might be a reasonable compromise between the approach that allows each parameter its own value, which can be problematic if there is sparse data, and the approach that assumes the effects are all the same, which neglects any possible differences among the individual chemicals.

Choosing between these two extreme possibilities is often unnecessary. The existence of estimators that compromise or average across both approaches (Morris, 1983a), results in a model than can outperform either choice (Greenland, 1999). This has been termed as "borrowing strength from the ensemble" (Morris, 1983a), in that an imprecise estimate for any particular chemical will be bolstered by the information contained in the estimates for other related chemicals.

The discussion above introduces the concept of exchangeability. For the situation at hand, it amounts to the belief that given our lack of knowledge about carcinogens and our inability to make distinctions, the effects of certain occupational substances might be thought to be similar, or even the same, until proven otherwise. Stated another way, we might say that learning of the causal relation of lung cancer to, say, barium chromate would necessarily increase our suspicion that lung cancer might also be causally related to basic lead chromate, or any of the other chromate compounds.

Exchangeability is a common concept in statistics and epidemiology (Greenland and Robins, 1986; Greenland, 1998; Lindley, 2000). It is directly related to the fundamental assumption of independently and identically distributed observations in probability theory. In the present context, by invoking exchangeability we can imagine that the K parameters to be estimated arise from identical distributions; in other words, the K coefficients are generated from K independent random draws from a single prior distribution (Thomas et al., 1985; Greenland, 1992).

The compromise-estimator can be represented by the following approximation, where B_1 to B_5 are considered exchangeable,

$$B_i^* = \overline{B}\,(1-W_i) + B_i W_i \quad ,$$

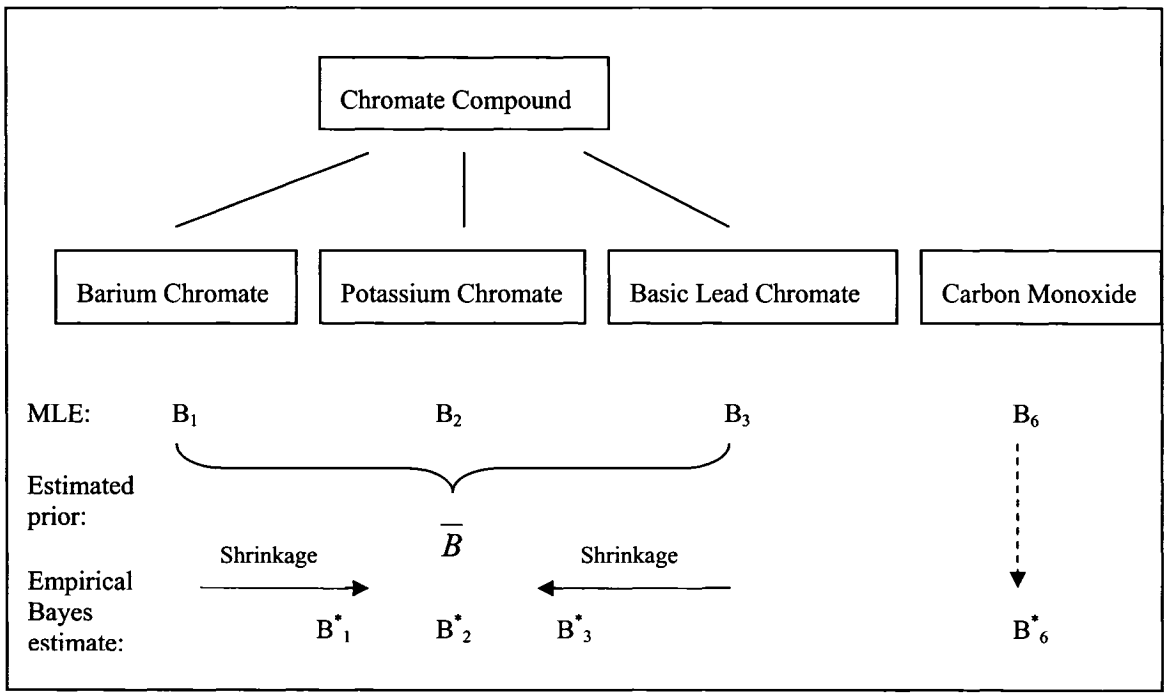
where W_i is some weighting function that takes into account both the variance of the individual B_i and the distribution of the ensemble of B_i. If the variance for B_i is large, implying little evidence is contained in that maximum likelihood estimate, then B_i^* will be shifted closer to the pooled estimate, the grand mean, $\overline{B}$. If the five B_i in our chromate example are widely distributed, implying B_1 to B_5 are not actually the same and that the assumption of exchangeability was questionable, then B_i^* will be shifted closer to the individual estimate, B_i. This is approximately the empirical Bayesian estimator. By specifying which parameters are thought to be similar (the exchangeability part), the prior mean is calculated as the average of the maximum likelihood estimates for these parameters (the empirical part), and then each MLE is in turn averaged with this prior mean (the Bayesian part). This can be viewed as a means of weighting different sources of information on the parameters of interest (Greenland, 1999).

To illustrate some of these concepts of related chemicals and exchangeable parameters, Figure 3-1 displays three of the five chromate chemicals mentioned above and a new chemical, carbon monoxide.

The principle behind this diagram is that of exchangeability. If no information exists to distinguish the effects of the three chromate compounds or to order the magnitudes of their effects, one must assume symmetry among the parameters in their prior distribution (Gelman et al., 1995). In other words, ignorance implies exchangeability, so that the less

that is known about the true effects, the more confident is the claim of exchangeable parameters. At the top of the diagram is one chemical characteristic, named "chromate compound." Three chemicals, barium chromate, potassium chromate, and basic lead chromate, share the common property of being chromate compounds. It is reasonable to assume that the individual effects of the three chromate compounds might be more similar to each other than they are to the effect of carbon monoxide. Specifying this in the context of the empirical Bayes model would result in the estimates of these three compounds shrinking toward each other. The B^*_6 of carbon monoxide is unaffected by this shrinkage. Thus goes the argument for exchangeability, and imposing such constraints on the model's estimation has several times demonstrated marked improvement over conventional analyses (Greenland, 1993).

Figure 3-1: Illustration of exchangeable parameters, using four occupational chemicals



The name empirical Bayes was first introduced by Robins (1955). Throughout the 1970s and 1980s these models were refined, culminating in the proposal for parametric empirical Bayes models (Morris, 1983a). Their introduction to mainstream epidemiology arguably dates to a publication by Thomas et al. (1985), with further developments occurring over the ensuing years (Louis, 1991; Greenland, 1992).

The basis for using these models has often been their ability to handle estimation in situations of sparse data, such as might occur if attempting to estimate regression models with large numbers of parameters in typically-sized epidemiologic studies. Bayesian models in general are not limited by the same rule-of-thumb restrictions of maximum likelihood estimation, that might otherwise lead to overfitting bias (Gelman et al., 1995); that is, the approximate 10:1 case to parameter ratio (Peduzzi et al., 1996). Although using large multi-parameter models is seemingly contrary to the parsimony principle, Savage has been attributed with the saying that "models should be as large as an elephant" (Greenland, 2000b). This antiparsimony viewpoint, at the very least, supports rich models for complex problems, and has been viewed favourably as a necessary counterbalance for those taking the principle of parsimony to an unnecessary extreme (Greenland, 2000b).

Many of the improvisations commonly used in model building, such as pre-testing of variables for inclusion into a model based on P-values or other criteria, and the deletion of near-collinear parameters (Holford, 2002), come at the expense of potentially introducing bias into the remaining parameters and often underestimating the model's residual variance (Freedman and Pee, 1989). The empirical Bayes approach is in essence an attempt to reduce parameters from the model, without actually introducing discontinuities in the modeling process (Greenland, 1992). This reduction is achieved by shrinking the parameter estimates closer toward the null, as opposed to setting them to the null value with absolute certainty by deleting them from the model.

These techniques have been variously described as a way of anticipating regression to the mean (Steenland et al., 2000), of recognizing chance occurrences (De Roos et al., 2001), of ensuring reduced overall error (Greenland, 2000a), of incorporating our knowledge in a quantitative manner (Robert, 2001), and of providing less arbitrary "best estimates" for ensembles of parameters (Thomas et al., 1985). The approach has been used in environmental (Morris, 1983a), occupational (Greenland, 1992; De Roos et al., 2001), biopharmaceutical (Louis, 1991), genetic (De Roos et al., 2003; Hung et al., 2004), dietary (Witte et al., 1994), geographic disease mapping (Marshall, 1991), and cancer surveillance (Steenland et al., 2000) studies, among others in epidemiology.

Empirical Bayes models are part of a family of hierarchical or multilevel models, which also include the following regression techniques as special cases: classical Bayesian, penalized likelihood, mixed models, ridge, and random-coefficients (Greenland, 2000a). The terminology found later in the methods section of the thesis will refer to the model with the parameters of interest as the first level, and the prior information as the second-level model.

3.3 Semi-Bayes models

Even though our knowledge may not be rich enough to specify the location of the prior (the mean), and hence the use of the EB estimator, it is likely that we can specify a range of uncertainty about the mean, especially if we believe that most strong carcinogens would have been previously identified (Greenland, 1992). In other words, we can often specify a range that will likely contain the true parameter values, and thereby "focus" the estimation. As opposed to the empirical Bayes method of estimating the prior mean and variance from the data, a step toward classical Bayes can be accomplished by assuming values for the prior variance and only estimating the mean from the data. This has been referred to as semi-Bayes modeling (Greenland, 1993).

Unlike empirical Bayes which requires iterative estimation, specifying the prior variance allows the equations to take on a simpler closed form (Greenland, 1992). As long as the true parameter values are, with reasonable certainty, to be found within the range specified, semi-Bayes methods will be an improvement over empirical Bayes. And as long as the specified prior variance is wide enough to assign a modest prior likelihood to the true value, semi-Bayes will outperform the usual maximum likelihood estimates (Greenland, 1993).

Specifying the prior variance requires an understanding of the hypothetical residual effects of the parameters of interest, after having taken into account the information contained in the second level of the model -- the exchangeability information (Witte et al., 1994). The range specified should reasonably encompass expert opinion about possible values of the parameters (Greenland, 1994). In a study of diet and breast cancer, Witte et al. (1994) specified nutrient levels that mediate the effects of food items. This involved prior information that the authors were quite certain would account for most of the effects

of the food items, and so a small, residual, twofold range for the prior standard deviation, T, was specified: exp(T*1.96*2)=2. On the other hand, in a study of occupational cancer, it could not be specified with much certainty which chemical and physical properties actually mediate carcinogenicity, and so allowance for a large prior standard deviation was necessary (De Roos et al., 2001): a tenfold range was specified, such that exp(T*1.96*2)=10. This states the belief that after having specified the sets of exchangeable parameters using chemical and physical properties of the occupational substances (similar to Figure 3-1), the residual effects of those substances should fall in a tenfold range centered close to unity, such as 0.3 to 3 on the ratio scale. By contrast, a conventional analysis, fitted into this framework, is tantamount to setting the prior variance to infinity and assuming that any empirical value is reasonable.

3.4 Performance of semi-Bayes estimators

Based on theory (Morris, 1983a; Greenland, 2000a), empirical Bayes estimators will provide overall gains in accuracy on an entire set of parameters, when compared to conventional maximum likelihood estimators, but any one of those estimates may be worse than its maximum likelihood counterpart. Morris emphasizes that for any particular parameter, the empirical-Bayes approach will provide these gains only insofar as one can assign a high prior probability to the true value of the parameter (Morris, 1983a).

These models require care in their design. Parameters that are clearly not exchangeable, need to be distinguished in the second level of the model (Greenland, 1992). Erroneously assuming that, say, tobacco smoke and wool fibres have exchangeable effects on lung cancer, will harm both estimates of their effects. Further, Greenland (1993) demonstrated that instead of allowing the prior variance to be estimated in empirical Bayes, gains in accuracy can be achieved with semi-Bayes modeling. Again, careful specification is necessary because under-specifying the prior variance can harm estimation (Greenland, 1993).

In simulation studies (Greenland, 1993; Witte and Greenland, 1996; Greenland, 1997), semi-Bayes results were compared to empirical Bayes, maximum likelihood, and pre-testing approaches in variously sized studies, some where asymptotic approaches should

be reasonable and some using small studies with many parameters. In small studies, empirical Bayes estimates showed improved accuracy and precision over maximum likelihood estimates. With few parameters and few study subjects, semi-Bayes models outperformed all other approaches, even with misspecification of the prior variance. In large studies with many parameters, EB and SB models performed similarly, with improvements over maximum likelihood and pre-testing approaches, but SB estimates were sensitive to any under-specification of the prior variance.

These simulations also demonstrated that there are restrictions to the number of parameters (representing exchangeability) that can be used in the second level of the model. If the model with the parameters of interest is referred to as the first-level model, the second-level model contains the covariates that define the categories of exchangeable first-level parameters. Additions of second-level parameters are at some point offset by the cost to precision, but Witte and Greenland (1996) nonetheless recommend that as many categories of exchangeability should be added as can be justified scientifically, as this should, in theory, still ensure an overall reduced mean squared error over the set of parameters being estimated.

4 The Montreal Study

4.1 Introduction

A large case-control study of multiple cancers was carried out in Montreal in the early 1980s by Siemiatycki (1991). This study served as the data source for the present thesis, and from here on this parent study will be referred to as the "Montreal study".

The purpose of the Montreal study was to investigate the possible relations between twenty-three different types of cancer and several hundred occupational substances and circumstances. This chapter gives a brief overview and the essential features of the Montreal Study, while the Methods section of chapter 6 will address those details that pertain specifically to the thesis.

4.2 Methods

4.2.1 Source population

Eligibility for the study was limited to men, aged 35 to 70, resident in metropolitan Montreal, and able to converse in French or English. Because cancer is a disease of long latency, and because the period of cancer occurrence was the early 1980s, the period of relevant exposure was roughly the 1940s through to the 1970s. In that era, relatively few women, as compared to men, worked in jobs involving heavy chemical exposures. Consequently, the cases of occupational cancer occurring in the early 1980s were more likely to be manifest among men than among women. The age restriction was also intended to focus on that subset of cases whose disease had the best chance of being attributable to occupational exposure. Cancer in younger or older men was thought to be more likely caused by other lifestyle factors or a purely genetic etiology.

4.2.2 Cancer series

Cases were ascertained through hospital pathology departments, requiring positive histological or autopsy evidence of a primary-occurring tumour occurring between the years 1979 to 1985. Participating hospitals were estimated to account for approximately 97% of Montreal's tumour diagnoses reported to the Quebec Tumour Registry (Siemiatycki, 1991). An introductory letter was delivered to the patient, and was

accompanied by a brief self-administered questionnaire, which inquired about certain demographic characteristics and lifetime work history. Subsequently, a face-to-face interview was conducted, which lasted on average 80 minutes, ranging from 5 to 120 minutes, depending on details of the work history. Some individuals refused to be interviewed but nonetheless accepted to complete a second, detailed self-administered questionnaire, which covered much of the interview material. Completed interviews were obtained on 3730 individuals with cancer.

The men with cancer were intended to be a source of cases and controls, whereby an analysis of any particular cancer site could use instances of the other cancers as the control series.

4.2.3 Electoral list series

A second series of men was recruited into the study by random selection from the elector list created for the 1981 Quebec provincial elections, using a probability sample matched to the age distribution of the cancer series, and restricted to metropolitan Montreal residents. After two years, difficulties arose due to many men having moved since the electoral list had been enumerated. At that time, random digit dialling was instead implemented for the period until the new electoral list of 1984 was enumerated. Completed interviews were obtained on 533 men.

The men recruited into the study from the electoral lists and the random digit dialling procedure were intended to be used as an alternative control series.

4.2.4 Data collection

Structured interviews were conducted to obtain information on important potential confounders, such as smoking history, age, education, income, ethnicity, alcohol use, diet, and hobbies. Semi-structured interviews were used for the occupational portion of the interview, allowing interviewers to probe for more information when it was warranted (Siemiatycki, 1984). When study subjects themselves were not able or available to be interviewed, a proxy respondent, like a spouse, was instead used. Study subjects were asked to supply detailed descriptions for all jobs held in their lives, and the interviewers attempted to draw out as much information as possible about the particulars of each job.

The interviewers used general questionnaires and a special set of questionnaires created by the chemists to help with more technical information. These questionnaires were tailored for particular jobs, such as bricklayers, carpenters, electricians, i.a., and information was elicited concerning the employer, the worker's environment, raw materials, final products, machinery, the presence of fumes or dusts, use of protective equipment, and other details that might provide clues as to incurred chemical or physical exposures.

4.2.5 Exposure assessment

The study incorporated an expert-based system of exposure assessment. A specially-trained team of chemists and industrial hygienists were provided with the detailed descriptions of each job in the subject's work history, and it was their responsibility to translate this information into codes for exposure to occupational substances. A list of 294 substances was compiled for the study. Since the primary purpose of the study was the discovery of unsuspected carcinogens, the list of exposures of interest was comprised mainly of substances for which there had previously been little or no research on human cancer risk. A wide variety of materials and chemical groupings were chosen. The list included well-defined chemicals, chemical groups, mixtures of known and somewhat fixed composition, mixtures of variable composition, complex materials, and, occasionally, general categories of substances. The main criterion for selection of these particular substances was their presumed 'high' prevalence in the Montreal area. Of the 294 substances that appeared on the original list, at least eleven chemicals are presently considered to have strong evidence for lung carcinogenicity in humans, and at least eight have evidence considered as suggestive (Siemiatycki et al., 2004).

Aside from the information from the interviewed subjects, the chemists made use of information from employers, from consultants with knowledge of particular industries, and from several bibliographic sources.

For each job, the chemists coded whether or not there had been exposure to any of the 294 substances. Following this, for each job-specific exposure to a substance, four aspects of the exposure were recorded:

1. Calendar years of exposure. The chemists recorded the year the exposure began and the year it ended.
2. Certainty of exposure. To reflect how confident the chemists were that a worker was exposed to a particular chemical, a 3-point ordinal scale was used, with 1 indicating possible exposure, 2 indicating probable exposure, and 3 indicating definite exposure.
3. Concentration of exposure. A 3-point ordinal scale was used to reflect the relative concentration of a chemical, with 1 indicating a relatively low concentration, and 3 indicating high concentration, such as from actual handling of the substance. The concentration of the chemical in the so-called general environment was used as a baseline, and exposure had to have exceeded this level in the occupation to be coded as present. As an aid in assigning these codes, the chemists made use of standard benchmarks, which were occupations that were known to frequently correspond to each of the levels. Low, medium, and high concentrations are not on an absolute scale and are, thus, not comparable between substances.
4. Frequency of exposure. The chemists used a 3-point ordinal scale to indicate the proportion of work time exposed to a chemical, with 1 indicating less than 5% of time and 3 indicating greater than 30% of time.

The chemists were blind as to any cancer diagnoses. One chemist would code each subject's file for exposures, and a second chemist would then review the original material and the first chemist's coding. Final exposure codes were decided by consensus, and over the ensuing years the chemists have reviewed their work repeatedly, revising the exposure codes.

4.3 Analytic strategies

Given the number of cancer sites and exposures in this database, and the system of expert exposure assessment, this data has over the years provided a basis for many analyses and studies. The database of the Montreal study consists of 4263 study subjects, distributed across many cancer sites. In particular, it has provided for the study of 294 different occupational substances, 98 occupational groups, and 77 types of industry. Several

publications have arisen from the Montreal study, some addressing methodological issues (Siemiatycki et al., 1988; Siemiatycki et al., 1989; Leffondré et al., 2002; Rachet et al., 2004) and some addressing, for example, particular groups of occupational substances (Siemiatycki et al., 1986; Siemiatycki et al., 1987) or particular sites of cancer (Siemiatycki et al., 1994; Parent et al., 1998; Parent et al., 2000). A monograph was also produced, dealing with this study in great detail (Siemiatycki, 1991).

5 Rationale and objectives of the thesis

5.1 Rationale

While tobacco smoke is clearly the most important risk factor for lung cancer, one reason occupational carcinogens are also of considerable importance is the involuntary nature of exposure and the prospect of avoidance or reduction of exposure in the workplace (Boffetta, 2004). Estimates for the fraction of lung cancer cases attributable to occupational exposures have varied greatly, ranging from below 1% to upwards of 40% (Simonato et al., 1988). Moreover, the importance of occupational exposures is greater in certain populations, such as blue-class, working men (Boffetta et al., 1995). The Montreal Study provides a basis for the examination of many previously unstudied occupational chemicals. While a few types of cancer in the study have been addressed in prior publications, a comprehensive assessment still remained for lung cancer.

Another reason supporting the importance of occupational cancer research is that, while understanding of the impact of occupational exposures on the risk of lung cancer has often lead to regulatory change and improved hygiene in the workplace, many so-called occupational chemicals spill over into the general environment. Engine exhausts have greatly contributed to urban air pollution, and while radon was originally studied only among uranium miners, it is also currently an active area of research because of the seepage of radioactive gas into homes (Boffetta, 2004). For these reasons, and a simple enumeration of all the chemicals currently used in the workplace for which there has been little health research, there exists the need to continue epidemiologic investigations into occupational lung cancer.

Given the number of exposures dealt with in the present thesis, addressing the potential for mutual confounding required special analytic methods. Empirical Bayes and semi-Bayes methods have been suggested as appropriate for occupational hazard surveillance (Greenland and Poole, 1994), though their applications have rarely been with a study of the dimensions of the Montreal study. The portion of the Montreal study's dataset, used for this research, involved 231 occupational exposures with complex, inter-related definitions, and so there was a need to explore different approaches to modeling the effects of the chemicals as well as methods of constructing the semi-Bayes models.

5.2 Objectives

The overarching objective of this thesis was to provide evidence for the independent effect of each of several occupational chemicals on the occurrence of lung cancer, both by using a method of expert exposure assessment, considered to be more accurate than approaches used in previous community-based studies, and by thoroughly taking into account confounding from other occupational substances, something not always explicitly considered. This led to the following substantive and methodological objectives:

Substantive objectives

1. To provide independent estimates for the effect for each of 231 occupational chemicals on the risk of lung cancer.
2. To select a subset of chemicals based on evidence of an effect on the risk of lung cancer, and follow this with other analyses aimed at exploring these relationships.
    a. To provide further evidence about the relationships of the selected chemicals to lung cancer, by exploring the influence of exposure duration, concentration, and timing on the risk of lung cancer.
    b. To provide independent estimates for the effect of each of the selected chemicals on the risk of particular histological subtypes of lung cancer.

Methodological objectives

3. To implement a semi-Bayes approach for analyzing a dataset with a large number of complexly related chemicals and correlated potential risk factors.
4. To compare results derived from the semi-Bayes analyses with those derived from more conventional regression methods, and to link any differences in the results to the formal properties of particular methods.

6 Methods

6.1 Data source

The present thesis used the database of the Montreal study (Siemiatycki, 1991). For a total of 4263 men, it included information that was used to construct variables on cancer diagnoses, demographic characteristics, non-occupational confounders, occupations, and occupational exposures.

6.2 Case series

The Montreal study attempted to enrol men diagnosed with cancer in any of the twenty largest hospitals in the Montreal area during the calendar period of September 1979 to June 1985. Not every case of every site of cancer was interviewed through all the years of the study. The case series for the present thesis was comprised of all histologically confirmed primary lung tumours -- topography ICD-9 codes 162.1 to 162.9 (American Medical Association, 1998) – admitted in the 2nd, 3rd, and 6th years of data collection.

Histological information was collected and of the 857 cases of lung cancer diagnosed in the study, the following major histological subtypes of lung cancer were identified: small cell carcinoma (159 cases), squamous cell carcinoma (359 cases), and adenocarcinoma (167 cases).

6.3 Control series

The Montreal study allows the choice between two types of control series. One strategy is to use cases of cancer sites, other than the lung, to form the control series. An alternative strategy is to use the men that were sampled from the electoral lists. Both strategies have their pros and cons, as discussed by Siemiatycki et al. (1981). In the monograph of the Montreal study, Siemiatycki (1991) preferred the strategy of using cancer patients with other types of cancer as the control series because it offered some coherence with the lung cancer series, it was larger than the electoral-based sample, and it minimized the chance of differential quality of information between the cases and controls. In the present analyses I opted for the same strategy. Nevertheless, the sensitivity of the results to this choice was addressed by replicating certain analyses using various options for the control series.

In using other cancer cases as the control series in our analyses, epidemiologic theory requires that the selected cancers are not themselves related to the exposures under study (Miettinen, 1985; Rothman and Greenland, 1998). For most of the occupational substances in the study, however, there is little evidence regarding their effects on cancer risk. This would make it difficult to justify the selection of only a few 'unrelated' cancers to form the control series. Moreover, one of the objectives of the thesis was to address mutual occupational confounding, and with many substances being assessed simultaneously, the difficulty of selecting cancer diagnoses unrelated to every one of the exposures became prohibitive. To avoid such difficulties, I decided to combine most of the other cancer sites into a single control series, hopefully "washing out" any outstanding relations to the exposures under study. Even if there were the occasional relationship between a particular exposure and a particular cancer site, I expect that such effects would be largely diluted. This is an arguably conservative approach as it could attenuate the empirical estimates of exposure effects on lung cancer. Any true effect would be expected to disappear, however, only if that substance increased the risk for all, or most, of the cancers by the same magnitude. Yet limited evidence suggests that most carcinogens have been found to not cause multiple cancers (Magee, 1978; Merletti et al., 1984).

Due to the study design, cancers with higher incidence rates were disproportionately represented in the control series. To minimize the possibility of giving too much weight to a particular cancer site that unbeknownst to me was related to a chemical under study, no individual cancer site was allowed to account for more than ten percent of the whole of the control series; this involved randomly removing sub-samples from certain cancer sites.

The following equation shows the theoretical impact on the RR if one of the constituent control sites was associated with a chemical being evaluated for its effect on lung cancer:

$$RR_{biased} = \frac{RR_{lung}}{\sum RR_{cancer}(n_{cancer} / n_{total})},$$

where RR_{biased} represents the biased rate ratio for lung cancer, RR_{lung} represents the true rate ratio for lung cancer, RR_{cancer} represents the true rate ratio for each cancer subtype,

and n_{cancer} represents the number of cases for each cancer subtype. For a plausible example of a large bias, the counts of each site of cancer (shown below in Table 7-1 of the results) can be substituted into the equation, and it can be assumed that the true rate ratio for the relation of lung cancer to one particular chemical is 2.0, and that the chemical was not related to any of the other cancers in the control series but for prostate cancer, with a rate ratio of 3.0. The resulting biased lung cancer estimate would have shifted from true RR_{lung}=2.0 to approximately RR_{biased}=1.7. It is likely that the more typical situation would have involved smaller magnitudes of effect and less bias. These errors are arguably important or ignorable depending on the context of the study, but here I believe them to be acceptable due to the nature of surveying so many chemicals.

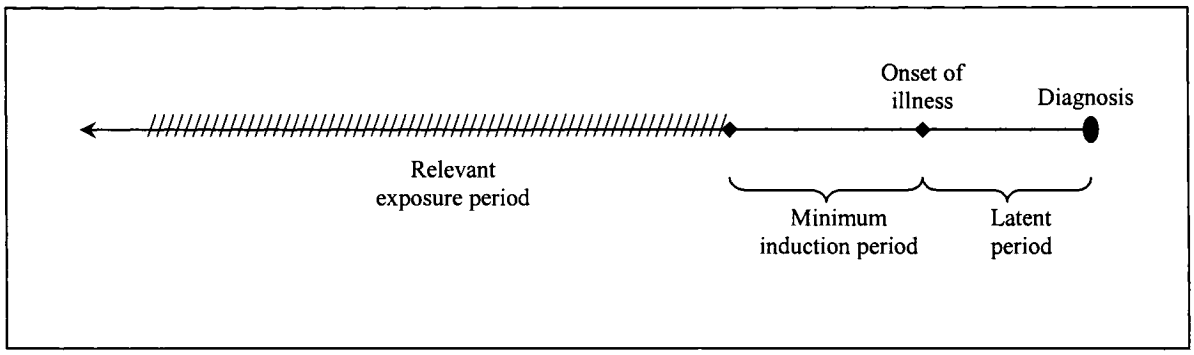
6.4 The occupational chemicals

From the initial list of 294 occupational substances, I chose not to address the extremely rare substances. Thus, exposures with a lifetime prevalence of less than 1% were removed from further analysis. An exception was made for a few rare chemicals that are currently suspected of being lung carcinogens (Siemiatycki et al., 2004): beryllium compounds, tobacco dust, jet fuel engine emissions, and cadmium compounds. The list of 231 chemicals assessed in the thesis can be found in the appendices.

6.5 Timing of exposures

In a study of etiology, the timing of exposures should be thought of as occurring on a temporal scale where the onset of illness is at t_0 of etiologic time (Miettinen, 1999). The moment of inception of a lung tumour is not in practice knowable, and so the year of diagnosis of the cancer was instead used. The difference between years of onset and diagnosis is often termed the latent period (Rothman and Greenland, 1998). It poses some difficulty in the attempt to define the period prior to diagnosis that is etiologically relevant for that diagnosis (Salvan et al., 1995), especially in the case of cancer which may require many years for the formation of a solid tumour. This difficulty is compounded by the problem of identifying the relevant minimum induction period for a particular exposure and lung cancer; that is, the time from causal action of an exposure to the initiation of the illness (Rothman and Greenland, 1998). Figure 6-1 illustrates these concepts on a time scale.

Figure 6-1: Induction and latent periods in the timing of exposure



In defining the period of irrelevant exposure I discounted any exposure that occurred in the five years prior to diagnosis, because it would be highly unlikely that exposure in this period could induce the observed case of lung cancer. It is a crude attempt to address the relevant risk period. For an actual occupational risk factor, if five years is less than the true sum of latent and induction periods, then its estimated effect would be somewhat attenuated.

No distinction was made between intermediate past and distant past exposures in the main body of results. The main analyses addressed the entire lifetime work history prior to diagnosis. The acceptability of this approach depends on how plausible it is for exposures very distant in the past to be etiologically relevant.

6.6 Representation of the exposures in the regression models

For each chemical that the chemists determined was present in a given job, the following four characteristics of exposure were taken into account: duration (in years), relative concentration (low, medium, or high), frequency (less than 5% of work period, 5 to 30%, or greater than 30%), and certainty (possibly, probably, or definitely present).

If exposure to a particular substance occurred in more than one job in an individual's work history, the corresponding exposure characteristics were combined into a weighted average, with weights proportional to the duration of each job. A value 0.5 was added to each job's duration to account for measurement error, for example, to avoid a zero duration for jobs beginning and ending in the same year. Duration of the exposure was then calculated as the sum of each job's duration, and it would end where the last relevant

job ended. Overlapping part-time jobs were handled so that their durations were not incorrectly summed.

For example, if an individual was diagnosed with lung cancer in 1980, and he was exposed to asbestos in four separate jobs in his work history, the relevant data would be represented as in Table 6-1.

Table 6-1: A hypothetical example of a worker with exposure to a given chemical in 4 different jobs

Job	*First year of job*	*Last year of job*	*Duration (years)*	*Concentration*	*Frequency*	*Certainty*
1	1953	1955	2.5	2 (medium)	2 (medium)	1 (possible)
2	1962	1972	10.5	2 (medium)	2 (medium)	3 (definite)
3	1973	1973	0.5	1 (low)	1 (low)	3 (definite)
4	1974	1979	5.5	3 (high)	1 (low)	2 (probable)

Uncertain exposures were coded as 'possible' by the chemists. These exposures were accounted for separately; therefore, in the case described in Table 6-1, the first job would not be included in the time-weighting. For the three jobs that 'probably' or 'definitely' occurred, each duration would then have 0.5 added to it. With the denominator summing to 10.5+0.5+5.5 = 16.5, the respective weights for the three jobs would then be,

10.5/16.5=0.636, 0.5/16.5=0.03, and 5.5/16.5=0.344

The new concentration, frequency, and certainty would be calculated as follows:

Concentration: 2*0.636 + 1*0.03 + 3*0.344 = 2.33

Frequency: 2*0.636 + 1*0.03 + 1*0.344 = 1.65

Certainty: 3*0.636 + 3*0.03 + 2*0.344 = 2.69

The new duration of the exposure to asbestos would be 16.5 years, and the newly assumed ending date would be 1979.

Each of the different characteristics of exposure -- concentration, frequency, certainty, and duration -- is of interest independently, but for simplicity and intelligibility, a single, ordinal, composite index of exposure with four levels was derived. The algorithm used to derive the overall exposure index is described in Table 6-2.

Table 6-2: Algorithm for defining composite exposure level

Exposure level	*Exposure category*	*Certainty of exposure*	*Concentration x Frequency* [a]	*Portion of duration more than 5 years prior to diagnosis or interview* [b]
0	Unexposed	-	-	-
1	Possible exposure	1 (possible)	any	any
1	Irrelevant exposure	Any	any	0 years
2	Moderate exposure	> 1 (probable, definite)	≤ 3	> 0 years
			> 3	>0 and <5 years
3	Substantial exposure	> 1 (probable, definite)	> 3	≥ 5 years

[a] The product of the concentration and frequency weighted-averages; [b] The portion of exposure duration that falls outside of the latent/induction period.

For a given chemical, unexposed individuals were assigned a level of 0. If the chemists were uncertain whether the individual had any cumulative lifetime exposure (certainty coded as 'possible'), no matter what the other characteristics, the assignment was with a level of 1. If all exposure occurred within the five-year latent/induction period, assignment was also with a level of 1. All other exposures were coded as either moderate or substantial, with substantial (level 3) requiring high concentration and frequency and at least five years cumulative exposure occurring more than five years prior to the latent/induction period.

When applied to the hypothetical data described in Table 6-1 for the individual who had a certainty of exposure greater than 1, a value for the product of concentration and frequency of greater than 3, and a large portion of the duration (16.5 - (5-(1980-1979)) =

12.5 years) outside the 5-year latent/induction period, the algorithm assigns an exposure level of 3, representing substantial exposure to asbestos. In this fashion, each individual was assigned a single lifetime-history exposure level for each of the 231 occupational chemicals.

Two levels of exposure were analyzed. Referring to Table 6-2, 'any' exposure during the lifetime work history corresponds to exposure levels 2 and 3; whereas 'substantial' exposure during the lifetime work history only corresponds to exposure level 3. In all cases, uncertain exposures (level 1) needed to be discriminated from the reference level of 'no exposure' in the regression models. Also, because of the overlap between 'any' and 'substantial' levels, two separate multivariable models were estimated for each substance:

$B_1X_{(\text{levels } 2,3)} + B_2X_{(\text{level } 1)}$, for 'any' level of exposure

$B_1X_{(\text{level } 3)} + B_2X_{(\text{levels } 1,2)}$, for 'substantial' level of exposure,

where B's are the estimated regression coefficients and X's are dichotomous variables.

In both cases, estimates of B_1, referring to the exposure level of interest, were reported in the tables, and estimates of B_2, referring to the extraneous uncertain exposures, were discarded.

6.7 Representation of the non-occupational confounders in the regression models

A large number of non-occupational characteristics were elicited in the original interviews (Siemiatycki, 1991). In the current analyses, the choice of which non-occupational confounders to include in the regression models was based exclusively on *a priori* considerations, rather than on data-driven criteria. Based on the published literature of the etiology of lung cancer, variables were included in regression models to represent the following set of eight potential confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

Age is routinely considered as an important potential confounder since it is strongly related to cancer risk and may be related to exposure history. Using ethnic origin as a potential confounder is motivated by its relation to genetic predisposition, diet, and other

social behaviours, all of which are difficult to conceptualize and document. Family income and years of education are often used as proxies for socio-economic status, which in itself is a vague concept, but is correlated with disease and possibly with putative risk factors for cancer (Mao et al., 2001). In the Montreal study, family income was not obtained from the interviews directly, but rather from the average annual income within the census tract of residence for each study subject, as ascertained from the Canadian census. Income and years of education may affect health via many mechanisms, including living conditions and the use of health services. Recreational activity was also included as a confounder because physical activity has been associated with cancer risk (Gotay, 2005). Alcohol consumption has recently been linked to a possibly small increase in risk for lung cancer (Bandera et al., 2001), and clearly smoking history would be an important confounder to include in any lung cancer analysis. Finally, a variable for respondent status was included because it was anticipated that proxy respondents, usually spouses, might contribute a different quality of information than self-respondents, and it is an attempt at controlling for this.

In order to find an optimal parameterization of the confounding variables, I conducted a small exercise using one of the occupational exposures as the prototype -- namely, silica. Both continuous and categorical variables were considered, and choices were made ad hoc by considering the Akaike Information Criterion (Akaike, 1974) and by taking into account how the different parameterizations affected silica's point estimate.

Table 6-3 lists the non-occupational confounders and the variables used to represent them in all the regression models. Cigarette-years and drink-years are calculated as the average daily number of cigarettes or drinks, respectively, multiplied by the years of use.

Adjustment for cigarette smoking presents difficult problems in conceptualization, documentation, and representation in a regression model. So few non-smokers with lung cancer existed in the database that restricting the study to non-smokers was not possible. The parameterization of the three variables described in Table 6-3 to represent cigarette smoking was originally suggested in a study aimed at estimating the effects of smoking, but this applies just as well to the adjustment for smoking as a confounder (Leffondré et al., 2002).

Table 6-3: Statistical variables representing non-occupational confounders

Confounder	*Operational definition for the analyses*
Cigarette smoking	W1=1 if ever smoked; W1=0 otherwise
	W2=natural log of cigarette-years [continuous, centered]
	W3=1 if quit smoking 2-5 years previously; W3=0 otherwise
	W4=1 if quit smoking 5-10 years previously, W4=0 otherwise
	W5=1 if quit smoking 10-15 years previously; W5=0 otherwise
	W6=1 if quit smoking >15 years previously; W6=0 otherwise
Age	W7=age in years [continuous]
	W8=age-squared [continuous]
Ethnicity	[reference is French]
	W9=1 if Jewish; W9=0 otherwise
	W10=1 if Italian; W10=0 otherwise
	W11=1 if Anglophone, European, or unspecified; W11=0 otherwise
Mean census tract income	W12=annual income [continuous]
Alcohol consumption	[reference is 0 drink-years]
	W13=1 if between 1 and 39 drink-years; W13=0 otherwise
	W14=1 if between 40 and 199 drink-years; W14=1 otherwise
	W15=1 if between 200 and 1199 drink-years; W15=0 otherwise
	W16=1 if between 1200 and 3070 drink-years; W16=0 otherwise
Respondent	[reference is self-respondent]
	W17=1 if proxy-respondent; W17=0 otherwise
Education	[reference is 0 to 5 years of education]
	W18=1 if 6 to 9 years; W18=0 otherwise
	W19=1 if 10 to 27 years; W19=0 otherwise
Recreational activity	[reference is rarely or never active in adult life]
	W20=1 if active once a week for six months; W20=0 otherwise

The variables indicating the years since quitting smoking were devised so that individuals that quit within the two years before the diagnosis, were still considered current smokers.

The logarithm of cigarette-years was used as it provided slightly better fit compared with a conventional cigarette-years parameterization (Rachet et al., 2004). Non-smokers were assigned a value of zero, and the values were centered among smokers so as to provide proper interpretation for the ever/never variable (Leffondré et al., 2002). That is, without centering, ever/never would be meaningless because among smokers, zero is not a possibility for cigarette-years, and among non-smokers, non-zero values for cigarette-years are impossible. Centering allows the rate ratio for the ever/never variable to be interpreted as the relative risk for a smoker with the average level of smoking compared to a never-smoker.

Data were occasionally missing on either age when quit smoking or current status of smoking. Simple imputation was used for such missing data. Among smokers, if data were missing for whether they had quit or not, it was assumed they were still smoking. For those having quit, if the age of quitting was missing, imputation was based on mean values from the remaining data. Table 6-4 lists imputed values for years since quitting smoking, conditional on age, alcohol consumption, and respondent status.

Table 6-4: Values for simple imputation of years since quitting smoking

		Alcohol consumption							
		≤80 drink-years				***> 80 drink-years***			
		Age <45	***Age 45-55***	***Age 55-65***	***Age > 65***	***Age <45***	***Age 45-55***	***Age 55-65***	***Age > 65***
Respondent	***Self***	1.7	3.3	5.6	6.2	0.2	1.8	3.4	4.6
	Proxy	1.6	1.5	3.6	3.7	1.9	0.4	1.2	1.4

6.8 Issues arising in regression modeling

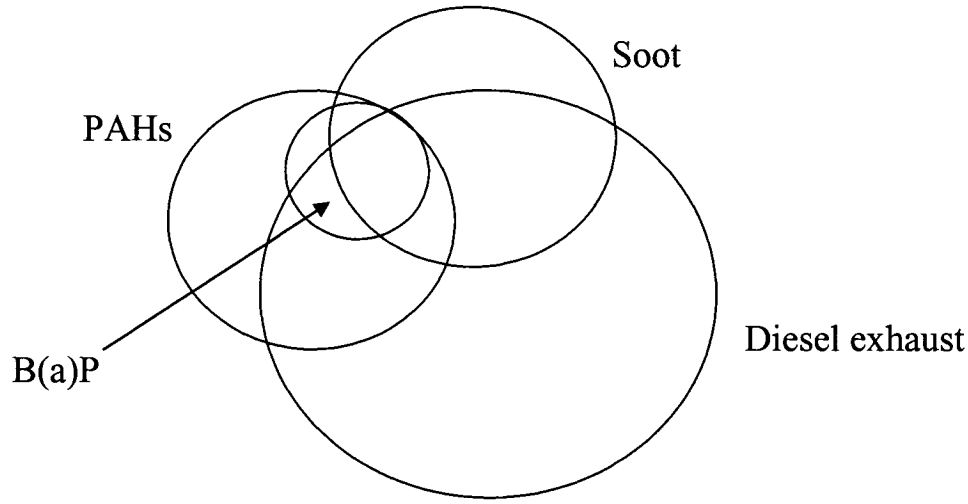
When considering the effect of a single occupational exposure, several issues arose in deciding which, if any, other occupational chemicals to include in the model as confounders. One issue is that the term chemical was used in this thesis to refer to both pure and complex substances. The list included complex mixtures of fixed or variable

composition, such as engine exhaust, which is composed of substances found elsewhere on the list. Chemical groups are also included, such as iron compounds (188), which encompass the presence of other specific substances on the list -- iron fumes (106), iron dust (33), and iron oxides (34). In a regression model, to adjust the general category for one of its sub-items would be to conceptually confuse the issues, induce artificial correlations between independent variables, and make interpretation of regression parameters difficult; for instance, if I included variables for both iron compounds and iron fumes in the model, I would have to interpret the parameter for iron compounds as the effect of increasing exposure to iron compounds other than iron fumes. Even more difficult would be the interpretation of the effect of iron fumes after adjusting for the effects of all iron compounds. This is also the case when adjusting complex chemicals, such as whole diesel engine exhaust (117), for its actual components, such as carbon monoxide (80) and nitrogen oxide (83); its parameter's interpretation would no longer be that for whole diesel exhaust.

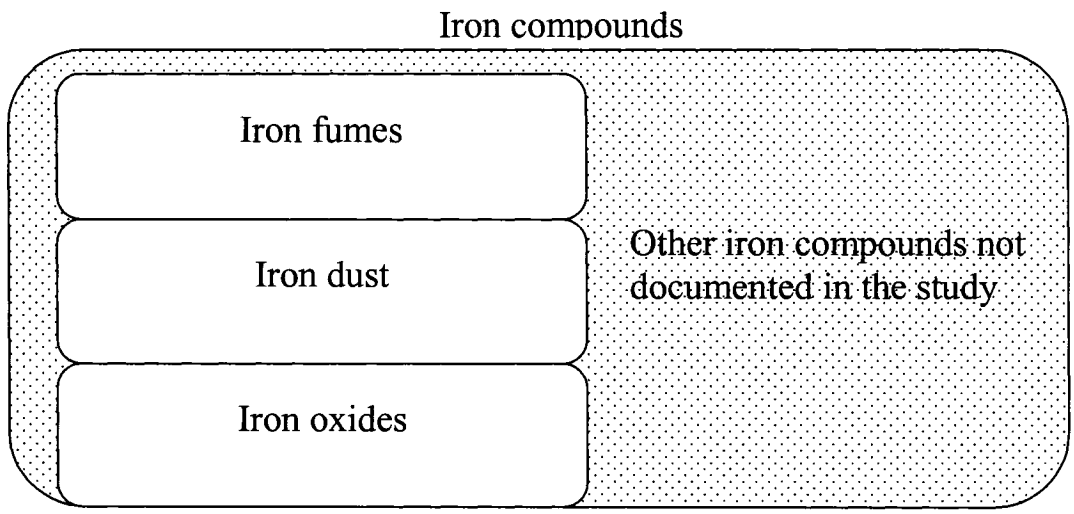
Figure 6-2 provides an illustration of two of the above difficulties in multiple regression modeling of determinants with complex relationships. Diagram A graphically describes that soot, a diesel engine emission, adsorbs polycyclic aromatic hydrocarbons (PAHs), which are also diesel engine emissions, and that elsewhere on the list is benzo(a)pyrene, which is a particular PAH. It is questionable whether to adjust the effects of these particular substances for each other. But given that each can occur from separate industrial processes in the workplace, they can also rightfully be considered potential mutual confounders. Diagram B graphically describes the general category of iron compounds, which was automatically coded whenever one of its constituents were coded, regardless of whether that constituent was found elsewhere on the list, or was not on the list, or was an iron exposure but determining which particular form of iron was difficult.

Figure 6-2: Illustration of complex relationships among occupational chemicals

A. Complex mixtures, illustrated with four chemicals found on the list of occupational exposures. Benzo(a)pyrene [B(a)P] is a PAH, which is a diesel engine emission and is adsorbed onto soot, which is also an engine emission. Each of these can occur from independent industrial processes in the population. The overlap of circles only indicates the conceptual relatedness of the definitions of these chemicals, and not correlation of exposures in the study population.



B. Chemical groups, illustrated with four chemicals found on the list of occupational exposures. Exposure to iron fumes, dusts, and oxides also automatically codes for exposure to the more general category of iron compounds. When the actual iron exposure could not be determined, 'iron compounds' alone was coded.



For the reasons just described, the regression models that included many occupational exposures simultaneously in the same model were analyzed by creating a restricted list of substances that could, at least in principle, have their effects be mutually adjusted without worrying about too many conceptual issues. From the list of 231 chemicals, a shorter list of 184 chemicals was created for analyses at 'any' level of exposure. Due to sparse data at the higher exposure levels, the list was further cut down to 146 chemicals for analyses at the substantial level of exposure. Table 6-5 lists those chemicals that were excluded from the larger models, along with the reasoning. I made a few exceptions to this rule, mostly because my interest in viewing their 'adjusted effects' outweighed the conceptual difficulties of interpreting the resulting parameter estimates. For example, I included gasoline engine emissions (115) and carbon monoxide (80) in the same model. For the purpose of testing how sensitive the results were to questionable adjustments, the results from a single large model, with nearly all 231 chemicals simultaneously analyzed, were also compared to the results from my preferred restricted model of 184 chemicals. As another sensitivity check, an even more restrictive model with only 117 chemicals was designed to avoid nearly every 'inappropriate adjustment' issue that was identified.

Only one instance of perfect collinearity was found in the dataset: PAHs from wood (216) and wood combustion products (119), but this was not a concern as neither was included in the larger models. On the other hand, many variables for the occupational exposures were highly correlated, some showing near-collinearity. Some sensitivity analyses avoided this issue by arbitrarily deleting one exposure from the model if a pair was found to have a Spearman correlation coefficient greater than 0.7. However, this could introduce bias to the remaining parameter, and recent work has suggested that the use of ridge regression can handle near-collinearity without the need for deletion (Le Cessie and Houwelingen, 1992; Holford et al., 2000; Holford, 2002). The purpose of maintaining such highly correlated variables in a single model would be to address the instability caused by collinearity and to attempt to differentiate the individual effects of the collinear variables. The empirical Bayes estimator can be considered a more generalized form of the ridge regression estimator (Greenland, 2000b). Thus, the single model with 184 chemicals retained highly correlated exposure variables, as opposed to deleting them.

Table 6-5: Chemicals excluded from the single regression model for the simultaneous estimation of chemical effects

Reason for exclusion	*Chemicals*
A chemical group	Metallic dust (4), fabric dust (72), rubber dust (77), synthetic fibres (54), cutting fluids pre 1955 (170), cutting fluids post 1955 (171), cyanides (176), fluorides (177), chromium VI compounds (178), magnesium compounds (182), aluminium compounds (183), titanium compounds (184), chromium compounds (186), manganese compounds (187), iron compounds (188), nickel compounds (190), copper compounds (191), zinc compounds (192), silver compounds (194), tin compounds (196), lead compounds (201), alkanes C18+ (202), alkanes C1-C4 (203), alkanes C5-C17 (204), aliphatic alcohols (205), aliphatic aldehydes (206), chlorinated alkanes (207), unsaturated aliphatic hydrocarbons (208), aliphatic esters (210), aliphatic ketones (211), wood PAHs (216), petroleum PAHs (217), coal PAHs (218), monocyclic aromatic hydrocarbons (220), aromatic alcohols (221), aromatic amines (222), cleaning agents (225), pharmaceuticals (226), laboratory products (227)
Mixtures, with constituents already modeled	Natural gas (90), mineral spirits + benzene/toluene/xylene (169), javel water (128)
Perfect collinearity	Wood combustion products (119)
Nebulous definitions	Other pyrolysis fumes (113), other mineral oils (166), other paints and varnishes (172), other PAHs (215)

* Numbers following chemical names refer to the numbering system of chemicals; see Appendix 2.

6.9 Regression modeling

The primary goal of the thesis was to estimate the effects of each of 231 chemicals on the risk of lung cancer. This was done using multiple strategies. All the models used for these analyses were based on the unconditional logistic regression model with maximum likelihood estimation, estimated with PROC LOGISTIC in SAS (1999). Some of the results utilized a semi-Bayes estimator, which used the logistic regression estimates as a starting point. The SAS matrix language program published by Witte et al. (1998) was

used for all the semi-Bayes analyses. Prior information for the semi-Bayes models was specified in the form of matrices within standard spreadsheet software.

The logistic transform takes the form,

Logit R =A+**BX**+**CW**,

where R is the incidence rate of lung cancer, A is an intercept term, **X** is a vector of variables representing each of the occupational exposures, and **B** is the vector of regression coefficients to be estimated. The model also includes potential confounders, represented by regression coefficients, vector **C**, and a vector **W** of variables, which for most of the analyses represent cigarette smoking, age, ethnicity, income, alcohol consumption, respondent status, education, and recreational activity. The estimate of effect used here is the odds ratio, which from the design of the study is an estimate of the incidence density-type rate ratio (Miettinen, 1976a; Greenland, 1987), derived from an exponentiated regression coefficient, e^B.

A fixed number of 231 chemicals were assessed. Each was given a unique number which is kept consistent throughout the results and discussion. A description of each chemical is found in the appendices.

The various modeling strategies necessitated different assumptions and criteria. The choice of models reflected different outlooks on how extensive control of confounding should be and what considerations should be taken into account while determining the optimal dimensions of the model. These analyses began with a simple model, such as one with only one exposure at a time and adjustment only for the confounding of age, and extend through to complex models that mutually adjusted for all other occupational chemicals and involved a Bayesian prior. The main modeling strategies are listed below in Table 6-6, along with a description.

The approach for strategy 1 represents a typical analysis in an occupational cohort study (Checkoway et al., 2004), where adjustment of each chemical's effect was only for age. Strategy 2 represented a conventional approach to such a large set of exposures, whereby the occupational chemicals were analyzed one-at-a-time, each within a separate regression model. Each exposure effect was adjusted for the standard set of non-occupational

confounders listed in Table 6-3. Strategy 3 was similar to strategy 2, but now each chemical-specific regression model was also adjusted for the following seven suspected lung carcinogens: asbestos (5), crystalline silica (6), chromium VI compounds (178), arsenic compounds (193), benzo(a)pyrene (219), diesel engine exhaust (117), and any source of polycyclic aromatic hydrocarbons (214).

Several issues arose when considering which chemical effects could be adjusted for each other. For the reasons described above, in section 6.8, 184 of the 231 chemicals were singled out to form a smaller list of chemicals whose effects could be simultaneously estimated in a single model. The remaining strategies involved models that were based on this shorter list of 184 chemicals.

Strategies 4 and 5 can be considered intermediate approaches between the 'one-at-a time approach' of strategy 3 and the 'single regression approach' of strategy 6, where all chemical effects were estimated simultaneously. In strategy 4, each chemical was analyzed in a separate regression model that used automatic forward selection with a P-value ≤ 0.25 criterion for the entry of other chemicals. The criterion of 0.25 was chosen to include chemicals with estimates that were statistically significant or marginally non-significant. Because P-values for the other variables would vary depending on which 'main exposure' was already forced into the corresponding model, the selection of covariates might vary from model to model. Each regression model began with the non-occupational variables forced into the model along with the one chemical that was the focus of that model. Automatic forward selection then proceeded to add the chemical that had the smallest P-value conditional on the variables already included in the model.

Strategy 5 involved estimating a single model where the only chemicals to be included were those that met an automatic forward selection strategy using a P-value ≤ 0.25 as the entry criterion. Unlike strategy 4, where all chemical effects were estimated, because each was analyzed in a separate regression model, in strategy 5 only a single logistic model was fit and the effects of chemicals not included in this model were set to zero on account of being excluded.

Table 6-6: List of main regression modeling strategies

Model strategy	*Short label*	*Number of Regression Models*	*Number of Chemicals assessed*	*Occupational chemicals adjusted as confounders*	*Non-occupational confounders*	*Bayes model*	*Location of results*
1	Separate models, age-adjusted only	231	231	0	Age	No	Estimates not shown
2	Separate models	231	231	0	All [a]	No	Table 7-3
3	Separate models, with seven suspected lung carcinogens	231	231	7 suspected carcinogens [b]	All [a]	No	Table 7-11
4	Separate models, Forward selection	184	184	Between 0 and 183 (chosen by P≤ 0.25)	All [a]	No	Table 7-11
5	Single model, Forward selection	1	Between 0 and 184 (chosen by P≤ 0.25)	Depends on which chemicals were selected	All [a]	No	Estimates not shown
6	Single model	1	184	All	All [a]	No	Estimates not shown
7	Single model, semi-Bayes, common prior	1	184	All	All [a]	Yes	Table 7-11
8	Single model, semi-Bayes, categories of exchangeability	1	184	All	All [a]	Yes	Table 7-11

[a] age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status;

[b] asbestos, crystalline silica, chromium VI compounds, arsenic compounds, benzo(a)pyrene, diesel engine exhaust, and polycyclic aromatic hydrocarbons from any source.

Strategy 6 involved a single regression model with all chemicals included, regardless of their statistical significance. Because of the size of the model, most of these estimates were expected to be imprecise, and previous publications have cautioned against using similar models for interpretation (Witte et al., 1994). For this reason, these estimates of the effects of chemicals are not presented in any table. This model's role was primarily as an interim step to the semi-Bayes approach, described below.

A semi-Bayes strategy was considered over the use of traditional empirical Bayes models. This was because of the greater control that can be exerted over the modeling, in terms of specifying prior variances from a scientific basis (Greenland and Poole, 1994), and to potentially gain improved estimation accuracy, as suggested in previous simulation studies (Greenland, 1993; Witte and Greenland, 1996).

Strategies 7 and 8 were semi-Bayes models (Greenland, 1992), and they built on the single model of strategy 6 by incorporating prior information about the parameters being estimated. This information was employed as assumptions of exchangeability about the effects of specific subsets of individual chemicals, meaning that within a subset of chemicals it was believed that their effects on lung cancer would likely be similar. For example, two chromate compounds would believably have more similar effects on lung cancer to each other than either would to the effect of, say, asbestos. Section 6.10 provides more detail about the implementation of these models. In brief, strategy 7 was designed to be a naïve model that treated all of the occupational chemical parameters as exchangeable. It used modeling strategy 6 as its first-level model, and the resulting estimates for each chemical were shrunk to a common prior mean. The prior variance for each parameter needed to represent a range of values that would most likely include the true magnitude of effect of each chemical (Witte et al., 1994). The prior variance was set to 0.345, which corresponded to a tenfold range of plausible values for the effects being estimated, such as 0.3 to 3 on the rate ratio scale.

In comparison to strategy 7, the approach to strategy 8 treated only subsets of parameters as exchangeable, based on shared chemical and physical properties (Greenland and Poole, 1994). These properties were represented in a second-level model, which described the parameters (effects of the chemicals) in the first-level model. These subsets were not

necessarily mutually exclusive. Chemicals could appear in multiple subsets as long as they were characterized by those particular chemical or physical properties. Thus, chemical-specific estimates from model 6 were shrunk in a multidimensional fashion, taking into account whichever subsets the chemical belonged to. The prior variance for each parameter was set to 0.246, which corresponded to a sevenfold range of plausible values for the effects being estimated. It was a smaller value because more information was included in the second-level model than had been included in strategy 7. The idea behind this is that as more information is included in the second-level model, less uncertainty remains about the chemicals in the first-level model (Witte et al., 1994). The model from strategy 8 also included information to indicate which parameters represented occupational exposures already suspected of being carcinogenic to lung tissue.

All of the above mentioned modeling strategies involve numerous assumptions. Some assumptions pertain to regression modeling in general, and some are peculiar to the various confounder selection strategies. Further, the parameterization of variables and the single index of exposure also involve assumptions. For example, the exposure index was based on an algorithm that multiplied ordinal values of concentration and frequency, with a further weighting by exposure duration, all of which involve assumptions about the relative effects of these exposure characteristics.

Only chemical-specific estimates from certain strategies were presented in tables. Other strategies were included for comparison purposes only, and their estimates were provided in the methodology-oriented results section, where they were often treated in aggregate. To be somewhat lenient in deciding which results to comment on in the text and follow-up in secondary analyses, 90% confidence limits were preferred over 95% limits. As a second rationale, a small technical advantage with 90% limits is that they more closely approximate the exact limits than would be the case with the 95% counterparts (Rothman and Greenland, 1998).

6.10 Implementation of the semi-Bayes models

The semi-Bayes approach relied on a two-level model, where the first level involved a simultaneous estimation of the effects of many occupational chemicals, using the usual maximum likelihood fitting method of logistic regression (strategy 6); and the second

level was a weighted-least-squares linear regression that modeled the first-level parameters (Greenland, 1992; Greenland, 1993), while accounting for the heteroscedasticity in the variances of the first-level estimates. This is a two-level hierarchical model, where the second level would be thought of by Bayesians as the prior. To simplify the fitting, a method-of-moments approach was used (Kass and Steffey, 1989; Marshall, 1991; Greenland, 1992). Method-of-moments can be thought of as a non-iterative first-order approximation to a maximum likelihood approach for acquiring the moments of a distribution, such as the mean and variance. It has been used in previous applications of semi-Bayes models (Greenland, 1993; Witte et al., 1994; De Roos et al., 2001). Two assumptions in using this approach were: that the distribution of maximum likelihood estimates was approximately normal, and that the likelihood would overwhelm any departures from normality, resulting in a normal posterior distribution even in the case of violations (Greenland, 1992).

To fit this two-level model, the following four steps were necessary: (i) estimate the effects of all the chemicals using conventional logistic regression, based on the model in strategy 6, (ii) specify a second-level matrix for these effects, (iii) specify the second-level residual intercepts, and (iv) specify the second-level residual variances. I used the SAS matrix language program of Witte et al (1998), which was modified to account for the study design. The program is listed in the appendices, with annotations, and was originally designed to handle empirical Bayes models, as well as the semi-Bayes variant. The difference between empirical Bayes and semi-Bayes lies in step (iv), where in the former the prior variances are estimated from the data, and in the latter the prior variances are specified according to scientific evidence.

Details of the four consecutive steps follow.

Step i

As a first step, the single regression model (strategy 6 from Table 6-6) with all chemicals represented simultaneously was fit with PROC LOGISTIC in SAS (SAS, 1999). The model for analyzing the 'any exposure' level was, Logit R=**XB** + **WC**, where **X** was the vector of 184 exposure variables, **B** was the vector of 184 exposure parameters, **W** was the vector of 20 non-occupational covariates plus 125 covariates for uncertain exposure

levels (not all exposures had sufficient numbers to warrant adjustment for this level of exposure), and **C** was the vector of the 145 parameters. The resulting first-level estimates were inherently unstable; that is, implausible estimates were expected because of the low number of cases compared to the large number of parameters being estimated in the model.

Step ii

The second step required specifying the second-level matrix of prior information. In the naïve approach to semi-Bayes regression (modeling strategy 7 in Table 6-6), all the 184 parameters of interest were considered exchangeable. In other words, this model shrank each of the 184 chemical-specific rate ratio estimates, obtained in step (i), toward the geometric mean of all the estimates combined, $(\prod rr_i)^{1/184}$ (Steenland et al., 2000). The extent of the shrinkage for each estimate was determined by its variance and the specified prior variance, described below.

Alternatively, expert knowledge can be brought to bear on the problem (Greenland, 1992), as in modeling strategy 8. In collaboration with chemists and industrial hygienists, I identified subsets of chemicals that were believed to have exchangeable effects based on shared chemical and physical properties. Such properties are plausibly relevant for lung carcinogenesis. Creating these categories of exchangeability was an iterative process that required several meetings with the chemists, hygienists, and epidemiologists. The chemists were first tasked with identifying as many major axes of chemical properties (elements, functional groups, and organic/inorganic characteristics) and physical properties (dusts, fumes, vapours, etc.) as they could. These were eventually combined in meaningful ways that categorized the list of 231 chemicals into subsets of related chemicals. These subsets were not mutually exclusive.

The effects of the chemicals in a particular subset were believed to be more similar to each other than to the effects of chemicals not in the subset, and I had little *a priori* reason to suspect the magnitude of effect of any one chemical in a subset to be lower or higher than any other chemical's effect in that subset. The chemists began with over 70 possible subsets, which I will refer to as categories, and eventually refined the list to 30 by removing categories which would have only contained one chemical or which were

deemed inappropriate for exchangeability purposes. For example, a category like 'solvents,' which was based on how chemicals were used, was discarded because it would not have clear implications for carcinogenicity. Removing categories was done cautiously because any property that was not included would *de facto* be considered of no consequence (Witte et al., 1994).

Table 6-7 identifies the categories specified for use in the semi-Bayes model strategy 8. Each of the chemicals can belong to one, to several, or to none of the categories. Chemicals not included in a subset still had their parameter estimates shrunk toward some overall prior mean, but wouldn't benefit from the influence of other chemicals with similar properties.

Table 6-7: Second-level categories, based on chemical and physical properties

No.	*Chemical and physical property*	*Description*
1	Polypeptides	Polymers of peptides (chains of two or more amino acids).
2	Polysaccharides	Type of carbohydrate: a major class of naturally occurring organic compounds. Polysaccharides are mainly sugars, starches and cellulose.
3	Fibrous inorganic dusts	Inorganic in nature, these dusts are threadlike strands, usually pliable and capable of being spun into a yarn.
4	Silica containing compounds	Compounds containing at least one atom of silicon (Si).
5	Metal dusts (excluding oxides)	Pure metal dusts only, excludes metallic oxides and other metallic compounds.
6	Metal oxide dusts	Oxides of any metal, in dust form.
7	Metal oxide fumes	Main component of the fume produced when a metal is heated to melting point.
8	Heavy metal compounds	Compounds containing metals of higher molecular weights.
9	Monocyclic aromatic hydrocarbons	Substances containing one and only one aromatic ring, regardless of other functional groups.
10	Polycyclic aromatic hydrocarbons	Substances containing more than one aromatic ring, regardless of other functional groups.

No.	*Chemical and physical property*	*Description*
11	Engine emissions	Complex mixture of substances produced by an internal combustion engine.
12	Inorganic acid mists	Inorganic acids in aerosol form.
13	Resins and resin-containing compounds	Gum-like substances obtained from trees or manufactured synthetically.
14	Carbonaceous compounds	Carbon-rich substances.
15	Aliphatic alkanes (C5-C17)	Hydrocarbons containing between 5 and 17 carbon atoms per molecule.
16	Aliphatic alcohols	Aliphatic compounds containing at least one hydroxyl group.
17	Aliphatic chlorinated hydrocarbons	Aliphatic substances containing at least one chlorine atom.
18	Inorganic gases	Inorganic compounds that are gaseous at room temperature.
19	Organic gases (C1-C4)	Organic compounds that are gaseous at room temperature.
20	Inorganic salts	Ionic compounds formed between the anion of an acid and the cation of a base; often soluble in water.
21	Magnesium compounds	Metallic compounds containing magnesium.
22	Aluminium compounds	Metallic compounds containing aluminium.
23	Chromates	Metallic compounds containing hexavalent chromium.
24	Manganese compounds	Metallic compounds containing manganese.
25	Iron compounds	Metallic compounds containing iron.
26	Nickel compounds	Metallic compounds containing nickel.
27	Copper compounds	Metallic compounds containing copper.
28	Zinc compounds	Metallic compounds containing zinc.
29	Tin compounds	Metallic compounds containing tin.
30	Lead compounds	Metallic compounds containing lead.

To avoid inappropriate shrinkage, it would have been necessary to distinguish strong determinants of lung cancer from weaker ones in the second-level model. For example, if smoking history was included as part of the semi-Bayes modeling (in fact, it was treated as a fixed effect, and the maximum likelihood estimate was not part of the second-level modeling), it would be inappropriate to include it in a category with other known occupational carcinogens. The effect of smoking on lung cancer is typically so strong, that its semi-Bayes estimate would be inappropriately shifted downward and the estimates of the other chemicals would be inappropriately shifted upwards. Distinguishing these effects by separate categories of exchangeability would rectify the problem. The parameters of the non-occupational confounders, such as smoking and age, were not of interest, and so they were treated as fixed effects; that is, they were not included in the second-level modeling and their estimates were not shrunk. Further, most occupational substances in the Montreal study were expected to have relatively small effects, and so errors in specifying categories of exchangeability would be expected to have only minor consequences to the results.

With the analysis in strategy 8, the maximum likelihood estimates of the first level tended to not only be shrunk toward their overall common mean, but they were also shrunk toward the means of the categories the exposures belonged to.

These categories were represented in the second-level model by dichotomous covariates. The data used for the second-level estimation was designed as a matrix (the z-matrix), with 184 rows corresponding to the 184 chemicals in the analysis and 31 columns corresponding to the 30 categories and 1 intercept (described in step iii). Each chemical was scored with a '1' in the appropriate column if it belonged to that category of exchangeability and a '0' otherwise. The approach for semi-Bayes strategy 7 did not implement categories of exchangeability, so the z-matrix was simply reduced to an 'intercept' vector of 1's, which implies that all chemical effects were deemed exchangeable. For illustrative purposes, Table 6-8 presents an extract from the second-level matrix for strategy 8. Substances with zeroes across all columns were not deemed exchangeable with any other substances; these were left out of any of the subsets, but their estimates were still shrunk toward the overall prior mean.

Such a matrix, as represented in Table 6-8, represents an attempt at creating a 'data equivalent' of a prior belief, which is then weighted along with the study data (Greenland, 2006).

Table 6-8: Selected elements of the second-level model: Covariates

	Categories of exchangeability (second-level covariates) [b]			
Chemicals [a]	*Fibrous inorganic dusts*	*Si-containing compounds*	*Metal oxide dusts*	*Inorganic acid mists*
Alumina	0	0	1	0
Silica	0	1	0	0
Asbestos	1	1	0	0
Diesel exhaust	0	0	0	0
Glass fibres	1	1	0	0
Phosgene	0	0	0	0

[a] Selected from the 184 possible chemicals. [b] Selected from the 30 possible properties used as categories of exchangeability

Step iii

Step three involved the specification of the second-level residual intercepts. That is, after taking into account the second-level information I specified in step ii for strategy 8, it may still have been possible to capture some of the predictable residual effect of the chemicals. For most, since I knew little about their effects, they were left with a near-zero prior mean. For the chemicals with strong or suggestive previous evidence of their carcinogenicity, however, I specified intercepts that equalled the expected residual log rate-ratios. These were taken from a recent meta-analysis of published studies on lung cancer (Steenland et al., 1996). Specifically, to be conservative, I used values close to the published lower 95% limits derived from those meta-analyses, since many of the studies focused on highly exposed workers. While all the rest were set to zero, the intercepts for the following chemicals were set to nonzero log rate ratio values (rate ratio in parentheses): silica, 0.18 (1.2); asbestos, 0.64 (1.9); iron oxides, 0.1 (1.1); diesel exhaust,

0.1 (1.1); arsenic, 1.1 (3.0); chromates, 0.8 (2.4); beryllium, 0.26 (1.3); nickel, 0.33 (1.4); cadmium, 0.18 (1.2); soot, 0.1 (1.1); coal gas, 0.1 (1.1); coal tar and pitch, 0.1 (1.1); polycyclic aromatic hydrocarbons, 0.1 (1.1); and sulphuric acid, 0.1 (1.1). Table 6-9 presents some of these elements, from the second-level model.

Step iv

Step four required specifying values for the second-level variance for each of the 184 parameters of interest. The "prior" variance is used as a scaling factor (Greenland, 1992), in that the amount of shrinkage of the logistic regression maximum likelihood estimate is in part determined by what was specified for the second-level variance. Based on prior knowledge of the magnitude of effects commonly seen in occupational lung carcinogenesis (Steenland et al., 1996), and with some knowledge of the exposure levels in the Montreal industrial environment (Ramzan Lakhani, personal communication), I surmised that most if not all of the effects being estimated in the study would manifest at somewhat low magnitude. Thus, I specified with 95% certainty that the true effects of the chemicals would most likely fall in a tenfold range about zero for rate ratios (tenfold implying, for example, 0.5 to 5). It was preferable to err on the conservative side by choosing values for the second-level variance that are large enough to encompass all reasonable opinions about the potential effects (Witte and Greenland, 1996). The value for this variance is calculated from a simple algebraic formula that represents with 95% certainty (1.96 standard deviations, much like 95% confidence intervals) that the residual effects will fall in a Q-fold range for rate ratios: $T^2=(\ln(Q)/3.92)^2$. For the naïve, intercept-only approach (modeling strategy 7 in Table 6-6), I specified a tenfold range, which corresponded to a prior variance of $T^2=(\ln(10)/3.92)^2=0.345$.

With the addition of the information on chemical properties and the information on previous evidence of carcinogenicity, I needed to envision what would be the residual effects of the first-level chemicals after having regressed out the 'effects' contained in the exchangeability information modeled in the second level. If everything important to occupational carcinogenesis was included in the second level, then the prior variances should be set to zero (Witte et al., 1994). On the other hand, with no information in the prior, the prior variance would effectively be infinity, and the maximum likelihood

estimates (from modeling strategy 6) would result. The "residual effect" is the effect theoretically left to the chemical after having accounted for any portion that might be attributed to the chemical and physical properties that were specified as second-level categories. For the approach with all the exchangeability information (modeling strategy 8), requiring in principle that there be less residual effect, I specified a sevenfold range (variance T^2=0.246).

Not really having any reason to distinguish which effects of the occupational substances was more or less "accounted for" by the chemical and physical properties, I set all the prior variances to the same value. Table 6-9 presents some of the elements in the second-level model for modeling strategy 8.

Table 6-9: Selected elements of the second level model: Residual effects

	Residual effects	
Chemical	*T^2 (range)*	*Intercept*
Alumina	0.246 (7)	0
Silica	0.246 (7)	0.18
Asbestos	0.246 (7)	0.64
Diesel exhaust	0.246 (7)	0.1
Glass fibres	0.246 (7)	0
Phosgene	0.246 (7)	0

The estimation

The final semi-Bayes model can be represented with these two levels:

Logit R = A + **XB** + **WC**, the first level, and

B=**ZP**+D, the second level, where **Z** is the vector of 30 second-level covariates (in the case of modeling strategy 8) plus 1 intercept, **P** is the vector of second-level parameters, and D is a random variable with mean zero and variance set to accord with the tenfold or sevenfold ranges mentioned above.

The posterior estimates were attained by averaging the estimated prior mean vector, **zp**, and the vector of maximum likelihood estimates, **b**, with weights proportional to the covariance matrix of the first-level estimates and the specified second-level variance, T^2 (Greenland, 1992).

Presentation of second-level parameter estimates

To appreciate the influence of the second-level covariates on lung cancer, the parameter estimates from the second-level model were also presented. In this context, the estimates would indicate the expected effect of the second-level characteristic on the risk of lung cancer (Witte et al., 1994).

6.11 Ranking and selection of chemicals

Although accurate estimation is a foremost concern, it is often useful to consider ranking of different exposures to prioritize them for further study (Thomas, 1985). In order to reduce the 231 occupational substances to a shorter list, chemicals were selected based on the evidence from the analyses outlined in Table 6-6. The first step in the selection was to rank the chemicals based on the strength of evidence. Simple methods of ranking are problematic for several reasons: ranking based on point estimates takes no account of the evidence represented by the variance, ranking based on P-values would not accurately account for the magnitude of effect, and ranking on a lower confidence limit is arbitrary in its choice of alpha level, where different choices lead to different ranking (Thomas et al., 1985; Thomas, 1985). The empirical-Bayes estimator, however, allows for a uniform approach to this problem (Thomas et al., 1985). A desirable characteristic of the empirical Bayes estimator is that it pulls imprecise estimates from the tails of the distribution of all estimates back toward the centre of the distribution (Thomas, 1985). This, in effect, anticipates regression to the mean (Thomas, 1985; Casella, 1985; Steenland et al., 2000), in that large imprecise estimates are more likely to be overestimates than not. In the re-ordering of estimates, extreme estimates with large variance will be shrunk substantially, sometimes even leapfrogging some initially weaker estimates with smaller variances (Berger, 1983). In theory, this allows for a better selection of chemicals with large point estimates.

6.11.1 The ranking

I ranked the chemicals using results from two of the modeling strategies: strategy 8, the semi-Bayes approach that assessed 184 chemicals in a single model, using categories of exchangeability based on shared chemical and physical properties; and strategy 3, a conventional approach that analyzed each of the 231 chemicals in a separate regression model, with adjustment for seven currently recognized lung carcinogens. Since more chemicals were analyzed in the approach for strategy 3, there was the possibility that more and/or different chemicals would be included in the eventual selection.

I ranked the chemicals based on the semi-Bayes point estimates of strategy 8, and then conducted a separate ranking based on the point estimates of strategy 3. Estimates for any level of exposure were ranked separately from estimates for the substantial level of exposure. This approach provided four separate lists of ranked chemicals.

6.11.2 The selection

To create the shorter list of chemicals with the strongest supporting evidence, I applied the following scheme to each of the four ranked lists. Focusing on potential causal (as opposed to preventive) substances, only chemicals with a point estimate above 1.0 were considered, and the chemicals had to also meet at least one of the following criteria:

- Chemical was a currently suspected or recognized lung carcinogen, based on previous evidence, or
- Point estimate was statistically significant with P-value ≤ 0.1, or
- Point estimate fell in the upper tail of the distribution, regardless of whether it was statistically significant or not. This was the only criterion that was explicitly based on the ranking of the chemicals. Operationally, this entailed selecting all chemicals descending the list until I reached the smallest estimate that remained statistically significant with P-value ≤ 0.1.

The models in strategy 3 did not have the same properties as the model of strategy 8, in terms of the latter's shrinking of estimates according to magnitude and precision. Thus, when selecting chemicals from the two lists based on strategy 3, the last criterion was not applied.

6.12 Attributable number of exposed cases

To provide an alternative to the rate ratio estimates, the number of exposed cases attributable to each of the chemical exposures (attributable number, AN) was also calculated and presented. Taking into account the number of exposed cases in the study population, the AN provides a useful public health indicator of the impact of an exposure on a population (Rothman and Greenland, 1998). The formula used to calculate AN was derived according to the formula of Miettinen (1974),

$$AN = \left(\frac{RR - 1}{RR} \right) C_E ,$$

where C_E is the number of exposed cases at either any level of exposure or substantial level of exposure.

It must be borne in mind that the estimates for the AN were based on both any and substantial exposure levels, and that these analyses were carried out independently of each other. Therefore, inconsistencies might arise. For instance, even though it is entirely possible to empirically observe a higher AN at the substantial level of exposure than at any level of exposure, since any exposure subsumes substantial exposure, this result would be logically impossible. Any apparent contradictory results reflect the weighting of true exposure levels in the 'any exposure' category and the statistical imprecision of the estimates.

6.13 Secondary analyses on the selected chemicals

The secondary analyses, for the most part, focussed on adding further evidence for each of the chemicals selected by the ranking and selection methods. Some of these analyses were undertaken to add more evidence as to whether or not there was truly a relationship with lung cancer. Other results may be used as further information about the characteristics of the exposure-lung cancer relationship.

6.13.1 Analyses of exposure characteristics

All the analyses listed above used an algorithm that combined exposure characteristics of concentration, frequency, certainty, and duration into a cumulative lifetime index. Addressing some of these characteristics independently could further inform the weight of

evidence for each chemical by providing insight into the form of the relationship with lung cancer. These analyses were restricted to the chemicals that were earmarked by the ranking and selection methods of section 6.11.

Three sets of analyses were aimed at exploring exposure characteristics. For each analysis, a separate regression model was fit for each chemical. For each chemical being in turn analyzed, men were deleted from the analysis if their exposure had been coded by the chemists as low certainty (certainty=1, see section 4.2.5). The operational definitions of the variables for these analyses are described in Table 6-10. Analyses included:

1. Analyses of concentration and frequency: Two levels of exposure were defined, combining the chemists' codes for concentration and frequency. Low to moderate levels of exposure, defined by having either characteristic coded as low or medium (1 or 2 on the 3-point ordinal scale), were differentiated from high levels of exposure, defined by having both characteristics coded as high (3 on 3-point ordinal scale). All results were adjusted for the eight standard non-occupational confounders.

2. Analyses of duration: Two exposure durations were modeled separately. In the analysis of short duration exposures, men exposed for more than 10 years were removed from the analysis. In the analysis of long duration exposures, exposed men with 10 or less years of exposure were removed. All results were adjusted for the eight standard non-occupational confounders.

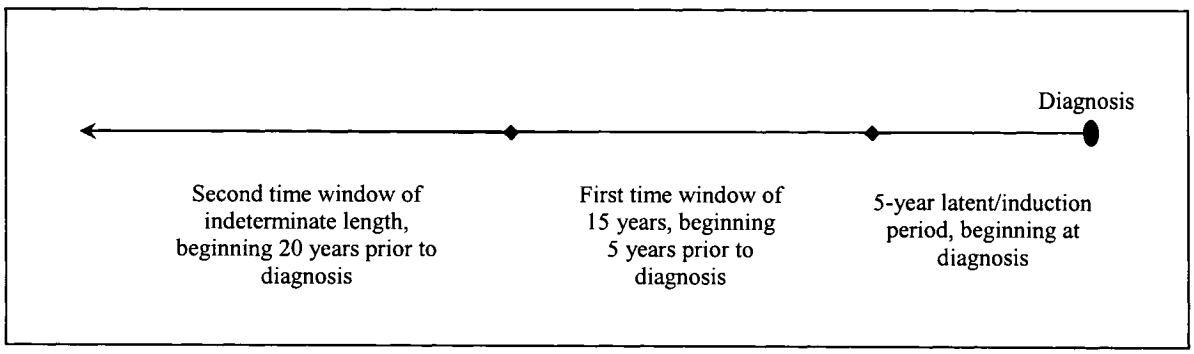
3. Analyses of time windows: In separate regression models, cumulative exposure within two separate windows of time were modeled, as illustrated in Figure 6-3. In the first regression model, exposed men were restricted to those that were exposed in the period 5 to 20 years before diagnosis. The same algorithm described in Table 6-2 for combining concentration, frequency, duration, and certainty, was used here, but now applied solely within this window of time. These analyses were based on any level of exposure within that window, and did not address the substantial level of exposure. The five years prior to diagnosis was again treated as an induction/latency period and exposures in this period were deleted from the analyses. In the second regression model, exposed men were

restricted to those that were exposed in the period 20 or more years prior to diagnosis. All results were adjusted for the eight standard non-occupational confounders.

Table 6-10: Statistical variables representing characteristics of exposure

Analysis	*Exposure characteristic*	*Operational definition of variables*
1	Concentration and Frequency	[reference is never-exposed] X1=1 if either concentration or frequency are low or medium (3-point ordinal scale values of 1 or 2); X1=0 otherwise X2=1 if concentration and frequency are both high (3-point ordinal scale values above 2); X2=0 otherwise
2a	Duration, 1-10 years	[reference is never-exposed] X1=1 if exposed for 1 to 10 years; X1=0 otherwise.
2b	Duration, 10+ years	[reference is never-exposed] X1=1 if exposed for 10+ years; X1=0 otherwise.
3a	Time window, 5-20 years prior to diagnosis	[reference is never-exposed] X1=1 if exposure occurred in the period 5-20 years prior to diagnosis; X1=0 otherwise.
3b	Time window, 20+ years prior to diagnosis	[reference is never-exposed] X1=1 if exposure occurred more than 20 years prior to diagnosis; X1=0 otherwise.

Figure 6-3: Illustration of time windows of exposure



6.13.2 Analyses with histological subtypes of lung cancer

Limited evidence suggests that some lung carcinogens act on particular histological cell types and not on others (Churg, 1994). Thus, where numbers permitted, analyses were conducted to examine associations between chemicals and histological subtypes of lung cancer. I carried out three separate sets of analyses. The case series was restricted in turn to one of the three major forms of lung cancer in the study: small cell, squamous cell, and adenocarcinoma. For each histology-specific case series, the chemicals were assessed in separate regression models, with adjustment for the standard eight non-occupational confounders (modeling strategy 2, see Table 6-6). Analyses were restricted to any level of lifetime exposure to a chemical, and did not include an assessment of the substantial exposure level.

6.14 Evaluation of the models

6.14.1 Comparison of results from different modeling strategies

One of the objectives of the thesis was to compare the estimates from the semi-Bayes models to those from the other modeling strategies, and to further tie the resulting estimates to the respective properties of the models. While studies of simulated data allow an evaluation of how well models perform with respect to true parameter values, I was limited to the existing dataset. Estimates were descriptively compared across the different modeling strategies. To summarize how or if the estimates tended to differ, various statistics were employed: the mean, median, skew, and kurtosis of the logistic beta estimates for the set of parameters, as well as the mean of the estimated standard errors.

Spearman correlations were used to compare the distributions of the ranks of estimates from pairs of models. Marginal distributions of logistic beta estimates and P-values were presented. Common criticisms of the P-value metric were not ignored (Goodman, 1993; Lang et al., 1998; Goodman, 2001), but I nevertheless sought to place the estimates from the models within the context of conventional views of the strength of evidence: RR<1 and P-value <0.05 to signify evidence of a preventive effect, and RR>1 and P-value<0.05 to signify evidence of a causal effect. P-values between 0.05 and 0.15, on both sides of

the scale, were used to indicate marginal evidence, and P-values $\geq$0.15, no matter which side of the scale the estimate lay, were used to indicate little or no evidence of an effect.

To highlight particular issues, two-way cross tabulations and scatter plots of the estimates were used when comparing pairs of modeling strategies.

A comparison of the strategies was also based on how many resulting estimates were statistically significant with P-values < 0.05 and how these numbers compared with expected numbers. Considering the 184 chemicals at any level of exposure and the 146 chemicals at the substantial level of exposure, a simplistic calculation was used to derive the number of estimates expected to be elevated merely due to random processes. This calculation was based on the assumption of a global null hypothesis that all rate ratios equal unity, asymptotic frequentist principles of 95% confidence intervals, and perfect validity of all the models. Based on the assumptions, one would expect, for any exposure and substantial exposure, respectively, 0.05*184=9.2 and 0.05*146=7.3 estimates with P-values below 0.05.

Statistical tests

Due to the number of chemicals being analyzed, not every estimate for every chemical that appeared in the tables of results was further described in the text. Selections of results for some narration in the text were often based on estimates with P-values $\leq$0.1.

As the main results were based on the control series design using cancer patients, a simple statistical test was used to indicate when one of these estimates appreciably differed from the estimate using the electoral list control series. For this purpose, I defined 'appreciable' as a one standard deviation difference. The calculation, which was only used when presenting results from the model that adjusted each chemical's effect for the eight non-occupational confounders (strategy 2, see Table 6-6), was arrived at by:

$$(b_{cancer}-b_{electoral})/\sqrt{SE^2_{cancer} + SE^2_{electoral}} > 1.0.$$

6.14.2 Sensitivity analyses

A common evaluation of regression modeling is to check for influential data and for lack of fit, but the more important concern here is clearly related to sparse data (Greenland,

1992). That is, the amount of data seems inadequate for the number of parameters involved in my larger models. The semi-Bayes approach has been proposed as a means of addressing such issues (Greenland, 2000b).

The results of many of the sensitivity analyses described below were presented as scatter plots, which allow a visual representation of how the estimates from different modeling strategies tended to differ.

Semi-Bayes models

I evaluated how dependent the results were on my preferred analytic choices by changing certain assumptions and characteristics of the analyses. The sensitivity analyses included two checks on the semi-Bayes models: different specifications of the prior variance, and a different method of handling previous evidence of lung carcinogenicity in the second level of the model. Results were compared from different models where the prior variances, in turn, were set to values representing infinity (the maximum likelihood estimate), a tenfold range, a sevenfold range, a twofold range, and zero (corresponding to the empirical Bayes model, where the prior variances were estimated instead of specified). Previous evidence, in my preferred approach, was represented as a second-level continuous covariate (see step iii, section 6.10). Again, results were compared from different approaches, one model where the previous evidence was instead represented by a dichotomous covariate (a zero indicating no previous evidence, a one indicating previous evidence), and a second model where previous evidence was not included in the model at all.

Full model for strategy 6

There were also several different versions of what constituted the single, large regression model (modeling strategy 6, from Table 6-6). While my preferred model included 184 chemicals at 'any level of exposure', I also analyzed models with different numbers of chemicals: a) A model that avoided any pre-selection at all, and included nearly all 231 chemicals simultaneously; b) A model including only 117 chemicals, removing all those that I had identified as posing conceptual difficulties and even chemicals that I retained in my preferred model despite minor conceptual issues (see section 6.8); and, c) A model

with variables deleted to avoid issues of multiple near-collinearity (Spearman correlations greater than 0.7).

Issues related to possible over-adjustment of estimates

As suggested above, minor conceptual issues remained in the list of 184 chemicals used for the single regression model, strategy 6. To assess the extent of any over-adjustment caused by adjusting certain troublesome chemicals for each other, a series of models were constructed that focussed on one issue at a time. For example, the point estimate of crystalline silica was tracked as other silica-based substances were added to the model one at a time. Other issues addressed in a similar fashion were: glass dust and glass fibres, natural rubber and styrene-butadiene rubber, whole engine exhausts and individual engine emissions, various sources for polycyclic aromatic hydrocarbons, various natural gases, general dusts and fumes, and the highly correlated chromium fumes and nickel fumes.

Control series options

Another evaluation included a comparison of results using three different designs for the control series: the cancer series, the electoral list series, and an amalgamation of the two. For the amalgamated control series, the 533 men from the electoral list were pooled with 533 men randomly chosen from the cancer series. Actual point and interval estimates for the effects of the chemicals under these three designs are shown in the appendices.

Respondent status

The analyses described above involved inclusion of respondent status as a covariate in the regression models. As a sensitivity analysis, an alternative strategy would be to restrict the analyses to the population of self-responders, which presumably would increase the accuracy of interview information.

Issues related to correlation and confounding

This section explores how the estimate for a particular chemical changed with the gradual addition of other chemicals, up to the size of the full model strategy of 184 chemicals (strategy 6). While focusing on the estimate for a single chemical, the intention was to assess whether the addition of so many other variables to the model would produce either

obvious or unpredictable patterns. A few chemicals were chosen for illustrative purposes, although only those with greater than 100 exposed cases were considered. With the focus on one chemical, say asbestos, 183 separate logistic models, each with asbestos and one of the 183 other variables, were fit. This allowed a calculation of the percentage change in the unadjusted estimate for asbestos due to each of the other chemicals. The chemicals were then reordered so that the first to be added to the gradually increasing model would be the chemical which had caused the greatest percentage change on the unadjusted estimate of asbestos, that is, the chemical that appeared to be the strongest confounder of the effect of asbestos. The second to be added to the model had caused the second greatest percentage change on the unadjusted estimate for asbestos, and so on. All 183 other chemicals were gradually added, allowing me to track the estimate for asbestos across all 184 iterations of the model. In all cases, the eight standard non-occupational confounders were included.

7 Results

7.1 Response rates

To be eligible for the Montreal study, patients had to be diagnosed with cancer in the pathology departments of the eighteen participating hospitals during the years 1979 to 1985 (Siemiatycki, 1991). Of the 4576 eligible individuals with cancer identified during the accrual period, completed interviews or self-administered questionnaires were obtained on 3730 individuals, an average response rate of 81.5%. Of the respondents, 82% had a face-to-face interview, 10% had a telephone interview, and 8% completed a self-administered questionnaire. Proxy respondents, such as spouses, accounted for 20% of the completions.

Table 7-1 lists a selection of the main types of cancer included in the study, the corresponding number of eligible individuals, and their rates of participation. In the main analyses, for the reasons listed in section 6.3, cancer sites other than lung were used to comprise the control series. Some cancers, such as cancer of the testis, contributed little to the control series, while others, such as cancer of the colon, would have contributed a large portion and so were reduced in number. The last column in the table shows the final number of cases of each cancer sub-type in the dataset.

The final cancer control series was comprised of 2172 individuals. Among these, 57 had primary cancers at two sites.

The second series of men used as the control series in some of the analyses were randomly drawn from Montreal electoral lists and a random digit dialling procedure. Of the 740 people contacted in the Montreal study, 533 completed an interview, a 72% response rate (Siemiatycki, 1991). I refer to these as the electoral list control series.

Table 7-1: Response rates for selected types of cancer

Type of cancer	*Number Eligible for Montreal Study*	*Response Rate (%)*	*Number of cases used in present analyses*
Lung	1082	79.2	857
Esophagus	129	76.7	99
Stomach *	318	78.9	215
Small intestine	37	78.4	22
Colon *	607	81.9	216
Recto-Sigmoid *	285	81.8	215
Rectum	304	84.5	190
Liver	76	63.2	48
Gallbladder	42	71.4	30
Pancreas	164	70.7	116
Prostate *	557	80.6	214
Testis	34	76.5	27
Penis	15	66.7	10
Bladder *	617	78.4	216
Kidney	227	78.0	177
Skin melanoma	124	83.1	121
N.H. lymphoma	258	83.3	216
Hodgkin's lymphoma	59	91.5	54
Myeloma	27	85.2	23

* Identifies those cancer sites which were restricted to 10% of the size of the control series (see section 6.3)

7.2 Characteristics of the study population

Table 7-2 presents the distributions of selected non-occupational characteristics separately for the case series, cancer control series, and electoral list control series. Compared to the cancer controls, there were more smokers among the case series. Both were similar in most other respects; the individuals in the case series, however, appeared to have fewer years of education, have less sustained recreational activity in adulthood, and were more likely to be of French origin.

Table 7-2: Distribution of selected characteristics in the study population

Characteristic	*Case series (n=857)*	*Cancer Control series (n=2172)*	*Electoral list Control series (n=533)*
Age			
35-50	78 (9%)	332 (15%)	60 (11%)
50-55	108 (13%)	276 (13%)	67 (12%)
55-60	191 (22%)	385 (18%)	96 (18%)
60-70	447 (52%)	1055 (49%)	264 (50%)
70-75	33 (4%)	124 (6%)	46 (9%)
Ethnicity			
French	592 (69%)	1238 (57%)	342 (64%)
Jewish	13 (2%)	106 (5%)	14 (3%)
Italian	48 (6%)	182 (8%)	37 (7%)
Other	204 (23%)	646 (30%)	140 (26%)
1981 Census tract income			
< 16k	151 (18%)	265 (12%)	57 (11%)
16-30k	629 (73%)	1514 (70%)	407 (76%)
30-50k	70 (8%)	351 (16%)	67 (12%)
50-60k	5 (1%)	11 (1%)	0 (0%)
60k+	2 (1%)	31 (2%)	2 (0%)
Alcohol index (drink- years)			
0	311 (36%)	1038 (48%)	299 (56%)
1-40	53 (6%)	214 (10%)	55 (10%)
40-200	288 (34%)	583 (27%)	132 (25%)

Characteristic	*Case series (n=857)*	*Cancer Control series (n=2172)*	*Electoral list Control series (n=533)*
200-1200	185 (22%)	316 (14%)	46 (9%)
1200 +	20 (2%)	21 (1%)	1 (0%)
Education (years)			
0-5	179 (21%)	340 (15%)	78 (15%)
6-9	391 (46%)	797 (37%)	255 (48%)
10+	287 (33%)	1035 (48%)	200 (38%)
Ever smoked cigarettes			
No	13 (2%)	375 (17%)	105 (20%)
Yes	844 (98%)	1797 (83%)	428 (80%)
Cigarette-years			
1-800	155 (18%)	775 (36%)	182 (34%)
801-1200	222 (26%)	485 (22%)	121 (23%)
> 1200	467 (54%)	537 (25%)	125 (23%)
Years since quit smoking			
Still smoking/recently quit	697 (81%)	1215 (56%)	363 (68%)
2-5	22 (2%)	217 (10%)	64 (12%)
5-10	23 (3%)	103 (5%)	25 (5%)
10-15	53 (6%)	149 (7%)	41 (8%)
>15	49 (6%)	113 (5%)	40 (8%)
Respondent			
Self	605 (71%)	1706 (79%)	466 (87%)
Proxy	252 (29%)	466 (21%)	67 (13%)
Recreational activity			
No	604 (70%)	1320 (61%)	287 (54%)
Yes	253 (30%)	852 (39%)	246 (46%)

7.3 Basic results for occupational chemicals

This section presents results from a number of simple approaches to assessing the effects of the occupational chemicals. It should provide some understanding of the distributions of exposures, the unadjusted estimates, and the estimates adjusted for non-occupational confounders only. Later sections will provide an evaluation and comparison of the different modeling strategies, and will present estimates from the more complex models.

Table 7-3 presents the tabulated counts of exposed and unexposed individuals used for the present analyses, allowing the derivation of unadjusted estimates for each substance. The table also shows estimates from modeling strategy 2, which included adjustment for eight non-occupational confounders (see Table 6-3).

The exposures documented in the study occurred over a wide range of prevalence, from a minimum of approximately 0.5% of the population, for chemicals such as beryllium compounds, to a maximum of 63% for exposure to some form of polycyclic aromatic hydrocarbons (PAHs). The varying precision of confidence intervals in Table 7-3 reflects this wide range. The median prevalence was 4.3%, with an inter-quartile range from 2.3% to 9.5%.

All estimates were presented at two exposure levels, any level of cumulative lifetime exposure and a substantial level of cumulative lifetime exposure. A few exposures are missing estimates at the substantial level of exposure, reflecting insufficient numbers to calculate those estimates. Some interval estimates, such as for fluorocarbons (212) at the substantial exposure level, appear to have a lower confidence bound of 0; this is merely rounding error, though it does reflect large imprecision.

Substance labels in italics indicate that that substance was not used in the larger models for the conceptual reasons listed in section 6.8 and Table 6-5. Substances with an asterisk indicate that the corresponding estimates using the electoral list series as controls differed appreciably from the estimates using the cancer series as controls, the latter forming the basis of the main results. The criterion I used for "appreciable" was a difference between the point estimates that exceeded one standard error, using the calculation described in section 6.14.1.

Section 7.8.3 provides an evaluation of the different control series design options, and estimates for all chemicals using the electoral list series as controls can be found in the appendices.

The results in Table 7-3 reflect which chemicals were associated with a higher risk of lung cancer, though not necessarily indicating causality. The models only included adjustment for non-occupational confounders and not for any of the effects of other correlated chemicals, thus presumably leading to overestimation. At any and substantial levels of exposure, 63 and 38 substances, respectively, had statistically significant estimates with P-Value<0.1. The following 25 substances had systematically elevated estimates: excavation dust, crystalline silica, Portland cement, metallic dust, borates, alumina, aluminium compounds, zinc dust, zinc fumes, zinc compounds, wood dust, nitrogen oxides, gas welding fumes, metal oxide fumes, manganese fumes, manganese compounds, iron fumes, iron compounds, copper fumes, copper compounds, solvents, kerosene, heating oil, mineral spirits with benzene-toluene-xylene, and alkanes (C5-C17).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
1. Abrasives Dust	610	1666	237	487	1.2 (1.0, 1.4)	100	237	1.0 (0.8, 1.3)
2. Inorg.Insul.Dust	739	1921	112	228	1.1 (0.9, 1.4)	36	62	1.2 (0.8, 1.8)
3. Excavation Dust	746	1961	109	198	1.5 (1.2, 1.8)	69	99	1.7 (1.3, 2.3)
4. *Metallic Dust*	569	1590	276	562	1.3 (1.1, 1.6)	138	253	1.5 (1.2, 1.8)
5. Asbestos	657	1795	177	335	1.2 (1.0, 1.4)	34	47	1.7 (1.1, 2.6)
6. Crystalline Silica	607	1663	238	480	1.3 (1.1, 1.5)	81	130	1.7 (1.3, 2.3)
7. Portland Cement	773	2014	79	141	1.4 (1.0, 1.8)	51	80	1.5 (1.1, 2.2)
8. Glass Dust *	839	2144	18	24	2.0 (1.1, 3.5)	6	15	0.8 (0.3, 2.0)
9. Glass Fibres	790	2009	50	130	0.9 (0.7, 1.2)	10	24	0.9 (0.4, 1.7)
10. Industrial Talc	795	2042	35	94	0.9 (0.6, 1.2)	8	28	0.8 (0.4, 1.6)
11. Brick Dust *	820	2085	34	81	0.9 (0.6, 1.3)	10	23	1.0 (0.5, 1.9)
12. Clay Dust *	823	2118	28	43	1.9 (1.2, 3.0)	9	15	1.6 (0.7, 3.3)
13. Concrete Dust *	759	1971	97	192	1.2 (0.9, 1.5)	58	85	1.6 (1.2, 2.2)
14. Bronze Dust	845	2145	11	27	1.0 (0.5, 1.9)	4	16	0.8 (0.3, 2.0)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
15. Brass Dust	833	2131	24	40	1.6 (1.0, 2.6)	10	15	1.6 (0.8, 3.3)
16. Stainless Steel Dust	803	2082	51	85	1.6 (1.2, 2.2)	20	36	1.6 (1.0, 2.7)
17. Mild Steel Dust *	684	1831	169	330	1.3 (1.1, 1.6)	79	169	1.2 (0.9, 1.6)
18. Inorg.Pigments	760	1979	93	187	1.3 (1.0, 1.6)	21	32	1.6 (0.9, 2.7)
19. Mineral Wool Fibres	781	2029	61	121	1.1 (0.8, 1.5)	10	25	0.7 (0.4, 1.4)
20. Extenders	805	2048	49	122	1.0 (0.7, 1.3)	11	30	0.8 (0.4, 1.5)
21. Aluminium Alloy Dust *	783	2049	63	115	1.5 (1.1, 2.0)	25	57	1.2 (0.8, 1.9)
22. Ashes	821	2120	36	51	1.4 (0.9, 2.1)	21	26	1.4 (0.8, 2.3)
23. Cosmetic Talc	840	2132	13	27	1.3 (0.7, 2.4)	2	13	0.4 (0.1, 1.4)
24. Borates	843	2150	11	20	2.0 (1.0, 3.9)	4	4	4.7 (1.2, 17.6)
25. Sodium Carbonate	841	2138	15	33	1.2 (0.7, 2.1)			
26. Alumina	683	1849	160	299	1.3 (1.1, 1.6)	26	45	1.6 (1.0, 2.5)
27. Silicon Carbide	797	2040	51	117	1.1 (0.8, 1.5)	10	19	1.2 (0.6, 2.4)
28. Sulfur	845	2136	9	31	0.7 (0.4, 1.4)	3	6	1.6 (0.4, 5.7)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
29. Calcium Oxide	733	1893	69	144	1.1 (0.8, 1.4)	25	46	1.4 (0.9, 2.3)
30. Calcium Sulphate	749	1959	100	197	1.2 (0.9, 1.5)	58	94	1.4 (1.0, 1.9)
31. Calcium Carbonate *	807	2025	46	143	1.0 (0.7, 1.3)	14	24	1.4 (0.8, 2.7)
32. Titanium Dioxide	816	2089	38	80	1.1 (0.8, 1.6)	6	7	2.0 (0.7, 5.5)
33. Iron Dust	818	2082	37	87	1.1 (0.7, 1.6)	11	37	0.8 (0.4, 1.4)
34. Iron Oxides	753	1951	101	213	1.1 (0.9, 1.4)	30	66	0.9 (0.6, 1.4)
35. Copper Dust	807	2075	47	92	1.3 (0.9, 1.8)	14	17	2.7 (1.3, 5.4)
36. Zinc Dust	829	2122	26	45	1.6 (1.0, 2.6)	5	5	3.7 (1.1, 11.7)
37. Zinc Oxide	821	2104	34	64	1.2 (0.8, 1.8)	4	6	1.1 (0.4, 3.6)
38. Lead Oxides	832	2137	22	34	1.8 (1.1, 2.9)	8	9	2.2 (0.9, 5.4)
39. Basic Lead Carb.	827	2126	28	42	1.4 (0.9, 2.2)	3	4	1.7 (0.4, 6.9)
40. Lead Chromate	822	2106	35	62	1.1 (0.8, 1.7)	3	4	1.7 (0.4, 7.3)
41. Organic Dyes & Pig.	783	1986	70	173	1.0 (0.8, 1.3)	14	26	1.4 (0.7, 2.5)
42. Cotton Dust *	791	1969	65	202	0.9 (0.7, 1.2)	30	105	0.8 (0.5, 1.2)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
43. Wool Fibres *	815	2026	41	143	0.9 (0.6, 1.3)	22	91	0.7 (0.5, 1.1)
44. Wood Dust *	620	1705	227	445	1.2 (1.0, 1.4)	122	218	1.3 (1.0, 1.6)
45. Grain Dust	787	2008	60	154	0.9 (0.7, 1.2)	26	59	0.9 (0.6, 1.4)
46. Flour Dust	818	2094	36	74	1.0 (0.7, 1.4)	16	44	0.7 (0.4, 1.1)
47. Fur Dust	843	2133	14	38	1.2 (0.7, 2.1)	6	18	0.8 (0.4, 2.0)
48. Hair Dust	848	2147	9	24	0.9 (0.5, 1.9)	5	16	1.0 (0.4, 2.5)
49. Starch Dust	837	2139	14	26	1.4 (0.8, 2.6)	5	12	1.3 (0.5, 3.5)
50. Sugar Dust	841	2148	15	21	1.6 (0.9, 3.0)	6	9	1.2 (0.5, 3.3)
51. Leather Dust	836	2095	21	74	0.7 (0.5, 1.1)	10	32	0.8 (0.4, 1.5)
52. Tobacco Dust *	848	2156	9	16	1.0 (0.5, 2.1)	3	5	0.8 (0.2, 2.8)
53. Natural Rubber	810	2075	44	88	1.2 (0.8, 1.7)	3	22	0.3 (0.1, 1.0)
54. *Synthetic Fibres*	811	2018	45	148	0.9 (0.7, 1.2)	20	80	0.7 (0.5, 1.2)
55. Plastic Dust *	810	2036	43	120	0.9 (0.7, 1.3)	11	36	0.9 (0.5, 1.7)
56. Rayon Fibres	836	2108	20	57	0.9 (0.6, 1.5)	10	22	1.1 (0.6, 2.3)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
57. Acrylic Fibres *	837	2112	18	53	0.8 (0.5, 1.4)	9	21	1.0 (0.5, 2.0)
58. Polyester Fibres *	826	2072	31	97	1.0 (0.7, 1.5)	16	59	0.8 (0.5, 1.3)
59. Nylon Fibres	833	2102	23	61	1.1 (0.7, 1.7)	10	28	0.8 (0.4, 1.6)
60. Acetate Fibres	843	2131	12	38	0.8 (0.4, 1.5)	3	11	0.8 (0.2, 2.4)
61. Cellulose Nitrate *	836	2106	18	57	0.7 (0.4, 1.2)	9	27	0.7 (0.4, 1.4)
62. Polyvinyl Chloride *	845	2130	11	36	0.6 (0.3, 1.2)	3	6	1.1 (0.3, 3.8)
63. Polyvinyl Acetate	834	2104	21	64	0.7 (0.4, 1.0)			
64. Poly-Acrylates	827	2106	29	60	1.3 (0.8, 1.9)	8	14	1.4 (0.6, 3.1)
65. Alkyds *	815	2082	39	88	1.0 (0.7, 1.5)	8	21	0.9 (0.4, 2.0)
66. Epoxies *	844	2144	13	24	1.8 (1.0, 3.5)	6	7	2.5 (0.9, 6.6)
67. Phenol-Formald. *	806	2087	47	67	1.5 (1.0, 2.1)	5	13	0.8 (0.3, 2.0)
68. Urea-Formald. *	799	2083	50	79	1.5 (1.1, 2.1)	4	12	0.7 (0.3, 2.0)
69. Polyurethanes	838	2136	18	30	1.4 (0.8, 2.4)	6	10	1.8 (0.7, 4.5)
70. Styrene-Buta.Rubber	817	2074	38	90	0.9 (0.6, 1.3)	2	19	0.2 (0.1, 0.9)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
71. Polychloroprene	827	2105	30	63	1.1 (0.7, 1.6)	3	5	1.1 (0.3, 4.0)
72. *Fabric Dust* *	780	1950	77	219	0.9 (0.7, 1.2)	39	122	0.9 (0.7, 1.3)
73. Coal Dust *	791	2069	63	101	1.4 (1.0, 1.9)	31	48	1.5 (1.0, 2.4)
74. Carbon Black	803	2067	52	100	1.3 (0.9, 1.7)	5	16	0.7 (0.3, 1.8)
75. Cellulose	790	2044	58	117	1.1 (0.8, 1.5)	16	36	1.0 (0.6, 1.8)
76. Soot	764	2003	91	166	1.2 (0.9, 1.5)	26	35	1.6 (1.0, 2.6)
77. *Rubber Dust*	824	2085	31	81	0.9 (0.6, 1.3)	2	17	0.3 (0.1, 1.1)
78. Graphite Dust	846	2140	8	26	0.7 (0.3, 1.4)	1	5	0.4 (0.1, 2.7)
79. Hydrogen *	836	2124	19	46	0.9 (0.5, 1.4)	3	7	1.2 (0.3, 4.1)
80. Carbon Monoxide	367	1103	478	1037	1.2 (1.0, 1.4)	108	188	1.3 (1.0, 1.6)
81. Hydrogen Cyanide *	838	2127	14	36	0.9 (0.5, 1.6)	7	16	0.7 (0.3, 1.6)
82. Ammonia *	762	1918	86	233	0.9 (0.7, 1.2)	28	90	0.8 (0.5, 1.1)
83. Nitrogen Oxides *	610	1748	240	414	1.6 (1.3, 1.9)	36	53	1.6 (1.1, 2.4)
84. Ozone	787	2044	67	121	1.5 (1.1, 2.0)	8	18	1.2 (0.5, 2.5)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
85. Hydrogen Fluoride	813	2120	38	49	2.0 (1.3, 2.9)	4	5	1.3 (0.4, 4.2)
86. Sulphur Dioxide	702	1837	144	325	1.1 (0.9, 1.3)	10	25	0.9 (0.4, 1.7)
87. Hydrogen Sulphide	816	2070	37	98	1.0 (0.7, 1.5)	8	19	0.8 (0.4, 1.6)
88. Chlorine	839	2115	15	57	0.5 (0.3, 0.9)	10	22	0.8 (0.4, 1.6)
89. Hydrogen Chloride	776	1981	59	149	1.0 (0.8, 1.3)	26	60	1.1 (0.7, 1.6)
90. *Natural Gas*	826	2116	24	54	1.0 (0.6, 1.5)	8	5	3.2 (1.2, 8.6)
91. Methane	809	2073	41	96	1.0 (0.7, 1.4)	13	13	2.8 (1.3, 5.7)
92. Propane *	811	2084	39	76	1.2 (0.9, 1.8)	2	7	0.8 (0.2, 3.4)
93. Formaldehyde	674	1680	125	308	0.9 (0.7, 1.1)	23	66	0.8 (0.5, 1.3)
94. Acetylene	808	2083	47	83	1.6 (1.1, 2.2)	1	5	0.4 (0.1, 3.0)
95. Phosgene *	843	2136	11	33	0.8 (0.4, 1.4)			
96. Spray Gases	841	2128	15	40	1.0 (0.6, 1.7)	3	16	0.5 (0.2, 1.5)
97. Coal Gas	847	2147	8	24	0.6 (0.3, 1.2)	2	3	1.5 (0.3, 7.4)
98. Gas Welding Fumes *	733	1939	115	211	1.5 (1.2, 1.8)	50	85	1.5 (1.1, 2.2)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
99. Arc Welding Fumes	747	1932	107	224	1.2 (0.9, 1.5)	44	72	1.5 (1.1, 2.2)
100. Soldering Fumes	794	2023	55	137	1.1 (0.8, 1.5)	31	64	1.3 (0.9, 2.0)
101. Metal Oxide Fumes *	660	1767	190	388	1.3 (1.1, 1.5)	87	142	1.6 (1.3, 2.1)
102. Aluminium Fumes	832	2128	23	42	1.4 (0.9, 2.3)	8	15	1.1 (0.5, 2.5)
103. Calcium Oxide Fumes	790	2036	66	131	1.3 (1.0, 1.8)	27	47	1.4 (0.9, 2.2)
104. Chromium Fumes	808	2112	43	53	2.2 (1.5, 3.3)	8	12	2.1 (0.9, 4.8)
105. Manganese Fumes *	794	2071	60	97	1.6 (1.2, 2.2)	14	15	3.4 (1.6, 7.0)
106. Iron Fumes	759	1990	94	172	1.4 (1.1, 1.8)	52	80	1.7 (1.2, 2.3)
107. Nickel Fumes	809	2110	42	55	2.1 (1.4, 3.1)	8	12	2.1 (0.9, 4.8)
108. Copper Fumes	808	2108	47	59	2.2 (1.5, 3.1)	16	19	2.3 (1.2, 4.2)
109. Zinc Fumes	814	2104	39	62	1.6 (1.1, 2.3)	16	18	3.0 (1.6, 5.8)
110. Silver Fumes	841	2141	15	30	1.4 (0.8, 2.4)	7	14	1.6 (0.7, 3.6)
111. Tin Fumes	804	2086	49	82	1.6 (1.1, 2.2)	10	19	1.2 (0.6, 2.5)
112. Lead Fumes	813	2086	41	78	1.4 (1.0, 2.0)	7	20	0.9 (0.4, 1.9)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
113. *Other Pyrolysis Fumes*	680	1820	171	339	1.3 (1.1, 1.6)	75	148	1.2 (0.9, 1.5)
114. Cooking Fumes *	798	2032	57	135	0.8 (0.6, 1.1)	29	69	0.9 (0.6, 1.3)
115. Gas Eng.Emissions	465	1266	379	879	0.9 (0.8, 1.1)	292	635	1.0 (0.8, 1.1)
116. Coal Comb.Products	803	2084	51	84	1.4 (1.0, 2.0)	19	36	1.1 (0.7, 1.9)
117. Diesel Eng.Emissions *	675	1835	165	295	1.2 (1.0, 1.5)	81	144	1.3 (1.0, 1.7)
118. Liquid Fuel Comb.Prod.	783	2039	71	131	1.2 (0.9, 1.6)	27	63	1.0 (0.6, 1.5)
119. *Wood Comb.Products*	814	2072	40	88	1.0 (0.7, 1.4)	17	31	1.1 (0.7, 2.0)
120. Natural Gas Comb.Prod.	827	2098	23	71	0.7 (0.5, 1.1)	6	16	0.9 (0.4, 2.1)
121. Jet Fuel Eng.Emiss.	854	2147	3	22	0.4 (0.2, 1.3)			
122. Propane Eng.Emiss. *	823	2128	28	41	1.7 (1.1, 2.6)	15	28	1.2 (0.7, 2.1)
123. Plastics Pyrol.Prod.	837	2096	17	72	0.6 (0.4, 0.9)	11	37	0.6 (0.3, 1.1)
124. Rubber Pyrol.Prod.	837	2120	20	50	0.9 (0.6, 1.5)	9	23	0.8 (0.4, 1.6)
125. Propane Comb.Prod. *	821	2096	30	64	1.1 (0.7, 1.6)	13	17	1.7 (0.9, 3.3)
126. Inorg.Acid Solutions	713	1859	129	264	1.2 (1.0, 1.5)	34	86	1.0 (0.7, 1.5)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
127. Alkali, Caustic Solutions	773	2019	72	141	1.3 (1.0, 1.7)	23	41	1.6 (1.0, 2.6)
128. *Javel Water*	808	2043	44	118	0.7 (0.5, 1.0)	30	64	0.9 (0.6, 1.3)
129. Plating Solutions	846	2148	10	22	1.1 (0.5, 2.2)	4	8	1.3 (0.4, 3.9)
130. Nitric Acid *	845	2138	9	31	0.8 (0.4, 1.6)	3	15	0.5 (0.2, 1.5)
131. Phosphoric Acid	841	2137	14	26	1.7 (0.9, 3.1)	2	7	0.8 (0.2, 3.6)
132. Sulphuric Acid *	719	1841	90	206	1.0 (0.8, 1.3)	13	38	0.8 (0.5, 1.5)
133. Methanol *	809	2062	44	103	1.0 (0.7, 1.3)	16	31	1.4 (0.8, 2.5)
134. Ethanol *	842	2136	15	32	1.4 (0.8, 2.5)	2	7	1.1 (0.3, 4.7)
135. Ethylene Glycol	816	2080	38	86	0.9 (0.6, 1.2)			
136. Isopropanol	814	2086	40	83	1.1 (0.8, 1.6)	15	30	1.3 (0.7, 2.3)
137. Acetic Acid	828	2086	28	82	0.9 (0.6, 1.4)	8	25	1.2 (0.6, 2.4)
138. Carbon Tetrachloride	809	2060	36	96	1.0 (0.7, 1.4)	19	35	1.6 (0.9, 2.6)
139. Methylene Chloride	839	2121	17	46	0.9 (0.6, 1.6)	8	16	2.0 (0.9, 4.6)
140. 1,1,1.-Trichlorethane	835	2137	16	24	1.8 (1.0, 3.3)	9	16	1.9 (0.8, 4.2)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
141. Trichloroethylene	825	2100	25	60	1.3 (0.8, 2.0)	9	31	0.8 (0.4, 1.5)
142. Perchloroethylene *	843	2138	11	29	1.1 (0.6, 2.1)	6	19	1.0 (0.4, 2.2)
143. Acetone	829	2110	20	51	1.0 (0.6, 1.7)	11	22	1.2 (0.6, 2.4)
144. Benzene	683	1779	162	367	1.0 (0.8, 1.2)	36	93	0.9 (0.6, 1.3)
145. Toluene	728	1852	120	302	0.9 (0.8, 1.2)	32	99	0.8 (0.5, 1.1)
146. Xylene	748	1905	96	244	0.9 (0.8, 1.2)	13	32	1.1 (0.6, 2.1)
147. Styrene	847	2122	10	44	0.5 (0.3, 0.9)	6	28	0.4 (0.2, 0.9)
148. Phenol	842	2134	11	30	0.8 (0.4, 1.6)			
149. Animal & Vege.Glues *	822	2086	34	79	1.0 (0.7, 1.4)	13	29	0.8 (0.5, 1.5)
150. Turpentine	795	2054	58	112	1.2 (0.9, 1.6)	33	59	1.3 (0.9, 1.9)
151. Linseed Oil	799	2074	53	91	1.3 (1.0, 1.8)	6	9	1.3 (0.5, 3.3)
152. Synthetic Adhesives *	718	1826	133	313	1.0 (0.8, 1.2)	59	165	0.8 (0.6, 1.1)
153. Solvents	470	1326	375	809	1.2 (1.0, 1.4)	227	487	1.2 (1.0, 1.5)
154. Waxes, Polishes *	795	2038	56	121	0.9 (0.7, 1.3)	18	37	0.8 (0.5, 1.4)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
155. Leaded Gasoline	733	1924	122	237	1.1 (0.9, 1.4)	50	104	1.1 (0.8, 1.5)
156. Kerosene *	784	2067	69	94	1.6 (1.2, 2.2)	26	21	2.7 (1.6, 4.5)
157. Diesel Oil	810	2090	45	78	1.4 (1.0, 1.9)	18	32	1.6 (0.9, 2.7)
158. Heating Oil	802	2091	53	79	1.4 (1.0, 2.0)	24	24	2.2 (1.3, 3.7)
159. Mineral Spirits *	741	1921	110	240	1.2 (1.0, 1.5)	64	134	1.3 (1.0, 1.8)
160. Lubric.Oils & Greases *	554	1519	291	626	1.2 (1.0, 1.4)	84	155	1.3 (1.0, 1.7)
161. Cutting Fluids *	769	1994	85	166	1.3 (1.0, 1.7)	33	69	1.2 (0.8, 1.8)
162. Asphalt *	824	2101	30	63	0.9 (0.6, 1.3)	13	22	1.1 (0.6, 2.1)
163. Coal Tar and Pitch	830	2125	23	41	1.0 (0.7, 1.7)	12	16	1.3 (0.7, 2.6)
164. Creosote	847	2137	5	26	0.7 (0.3, 1.7)	2	9	0.8 (0.2, 3.3)
165. Hydraulic Fluid *	818	2086	37	81	1.0 (0.7, 1.4)	4	8	1.3 (0.4, 3.8)
166. *Other Mineral Oils* *	818	2086	32	78	1.0 (0.7, 1.4)	5	14	0.8 (0.3, 2.0)
167. Jet Fuel	851	2143	6	26	0.6 (0.3, 1.4)	4	18	0.6 (0.2, 1.6)
168. Aviation Gasoline	851	2143	6	28	0.6 (0.3, 1.2)	4	19	0.6 (0.2, 1.5)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
169. *Mineral Spirits+BTX*	687	1845	158	314	1.3 (1.0, 1.5)	100	196	1.3 (1.0, 1.6)
170. *Cutting Fluids pre 1955*	791	2042	65	127	1.3 (1.0, 1.7)	21	41	1.4 (0.9, 2.3)
171. *Cutting Fluids post 1955* *	793	2058	62	102	1.7 (1.2, 2.3)	25	46	1.5 (0.9, 2.3)
172. *Other Paints, Varnishes*	729	1879	124	277	1.1 (0.9, 1.4)	57	111	1.2 (0.9, 1.6)
173. Wood Varnishes, Stains	797	2060	56	104	1.2 (0.9, 1.7)	32	58	1.4 (0.9, 2.0)
174. Inks *	817	2100	37	69	1.5 (1.0, 2.2)	17	40	1.1 (0.7, 1.9)
175. Metal Coatings *	780	2010	74	147	1.2 (0.9, 1.6)	30	62	1.2 (0.8, 1.8)
176. *Cyanides* *	834	2118	17	43	1.0 (0.6, 1.7)	9	19	0.8 (0.4, 1.7)
177. *Fluorides*	809	2109	42	60	1.8 (1.3, 2.7)	6	10	1.4 (0.5, 3.6)
178. *Chromium (VI) Comp.* *	758	1980	90	163	1.4 (1.1, 1.8)	12	26	1.5 (0.8, 2.9)
179. Hypochlorites *	806	2036	45	124	0.7 (0.5, 1.0)	30	66	0.9 (0.6, 1.3)
180. Nitrates	848	2137	8	23	0.9 (0.4, 1.9)	2	7	0.6 (0.2, 2.4)
181. Beryllium Compounds	851	2161	5	9	1.1 (0.4, 3.0)			
182. *Magnesium Compounds*	837	2137	19	29	1.9 (1.1, 3.4)	6	11	2.2 (0.9, 5.8)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
183. *Aluminium Compounds* *	635	1767	199	370	1.4 (1.2, 1.7)	51	96	1.6 (1.1, 2.2)
184. *Titanium Compounds*	808	2068	44	96	1.1 (0.8, 1.6)	7	12	1.4 (0.6, 3.4)
185. Vanadium Compounds	841	2150	16	22	1.5 (0.8, 2.8)			
186. *Chromium Compounds* *	716	1901	130	242	1.4 (1.1, 1.8)	15	33	1.4 (0.8, 2.6)
187. *Manganese Compounds*	784	2049	71	120	1.5 (1.1, 2.0)	14	16	3.0 (1.5, 6.1)
188. *Iron Compounds* *	604	1651	248	505	1.2 (1.0, 1.5)	122	228	1.4 (1.1, 1.7)
189. Cobalt Compounds *	833	2129	20	40	1.4 (0.9, 2.3)	4	14	0.8 (0.3, 2.1)
190. *Nickel Compounds*	770	2029	79	133	1.7 (1.3, 2.2)	12	22	1.7 (0.9, 3.3)
191. *Copper Compounds* *	725	1917	128	241	1.3 (1.1, 1.6)	55	101	1.6 (1.1, 2.1)
192. *Zinc Compounds*	743	1957	107	200	1.4 (1.1, 1.7)	25	30	2.4 (1.5, 3.9)
193. Arsenic Compounds *	816	2087	31	74	0.8 (0.6, 1.2)	11	25	0.9 (0.5, 1.8)
194. *Silver Compounds* *	829	2112	24	58	1.2 (0.8, 1.9)	10	21	1.7 (0.9, 3.5)
195. Cadmium Compounds	845	2151	11	18	1.7 (0.9, 3.5)	4	7	1.6 (0.5, 4.8)
196. *Tin Compounds* *	758	1984	92	177	1.4 (1.1, 1.8)	14	28	1.3 (0.7, 2.3)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
197. Antimony Compounds *	831	2127	19	39	1.3 (0.8, 2.2)	4	10	1.1 (0.4, 3.2)
198. Tungsten Compounds *	846	2141	11	30	1.4 (0.7, 2.6)	5	15	1.0 (0.4, 2.4)
199. Gold Compounds	842	2148	12	20	2.0 (1.0, 4.0)	4	7	2.6 (0.8, 8.4)
200. Mercury Compounds	836	2122	17	35	1.3 (0.7, 2.2)			
201. *Lead Compounds*	406	1153	434	979	1.0 (0.9, 1.2)	55	99	1.1 (0.8, 1.6)
202. *Alkanes (C18+)*	518	1439	320	697	1.2 (1.0, 1.4)	101	185	1.3 (1.0, 1.7)
203. *Alkanes (C1-C4)* *	767	1978	82	188	1.1 (0.8, 1.4)	15	23	1.8 (1.0, 3.3)
204. *Alkanes (C5-C17)*	479	1429	368	717	1.4 (1.2, 1.7)	207	376	1.5 (1.3, 1.8)
205. *Aliphatic Alcohols*	768	1969	83	191	1.0 (0.8, 1.3)	34	75	1.1 (0.8, 1.6)
206. *Aliphatic Aldehydes*	643	1586	146	378	0.8 (0.7, 1.0)	28	83	0.8 (0.5, 1.2)
207. *Chlorinated Alkanes* *	751	1908	93	210	1.1 (0.9, 1.4)	41	83	1.5 (1.0, 2.1)
208. *Unsat.Aliph.Hydrocarb.*	800	2060	54	105	1.5 (1.1, 2.0)	1	6	0.3 (0.0, 2.1)
209. Chlorinated Alkenes	793	2018	42	103	1.2 (0.8, 1.7)	15	46	1.0 (0.6, 1.6)
210. *Aliphatic Esters*	826	2093	29	70	1.1 (0.7, 1.6)	16	35	1.3 (0.7, 2.2)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
211. *Aliphatic Ketones*	795	2006	47	134	0.8 (0.6, 1.1)	22	63	0.8 (0.5, 1.3)
212. Fluorocarbons	839	2114	16	56	0.6 (0.4, 1.1)	1	18	0.2 (0.0, 1.1)
213. Glycol Ethers	827	2115	26	50	1.1 (0.7, 1.7)	8	6	3.7 (1.4, 9.9)
214. PAH (Any) *	230	749	581	1327	1.1 (1.0, 1.4)	80	130	1.4 (1.0, 1.8)
215. *PAH (Other)*	662	1755	187	402	1.2 (1.0, 1.4)	80	182	1.0 (0.8, 1.3)
216. *PAH (Wood)*	814	2072	40	88	1.0 (0.7, 1.4)	17	31	1.1 (0.7, 2.0)
217. *PAH (Petroleum)* *	269	851	561	1275	1.1 (1.0, 1.3)	393	828	1.2 (1.0, 1.4)
218. *PAH (Coal)*	766	2013	84	146	1.4 (1.0, 1.8)	35	62	1.2 (0.8, 1.8)
219. Benzo(a)pyrene	622	1703	220	421	1.2 (1.0, 1.4)	42	71	1.3 (0.9, 1.8)
220. *MAH*	513	1423	331	710	1.2 (1.0, 1.4)	97	203	1.2 (0.9, 1.5)
221. *Aromatic Alcohols* *	809	2039	21	69	0.9 (0.6, 1.4)	2	23	0.3 (0.1, 1.0)
222. *Aromatic Amines* *	798	2006	55	147	0.9 (0.7, 1.2)	3	20	0.4 (0.1, 1.2)
223. Phthalates	836	2101	15	56	0.5 (0.3, 0.9)	4	8	1.2 (0.4, 3.8)
224. Isocyanates	840	2136	16	33	1.0 (0.6, 1.7)	2	5	0.9 (0.2, 4.0)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

Table 7-3: Exposure frequencies and rate ratio estimates, from strategy 2 [a], for 231 chemicals, at two levels of exposure

	Unexposed		Any exposure			Substantial exposure		
	Cases	Controls	Cases	Controls	RR (90% CL)	Cases	Controls	RR (90% CL)
225. *Cleaning Agents* *	695	1811	154	335	1.0 (0.8, 1.2)	89	208	0.9 (0.7, 1.1)
226. *Pharmaceuticals*	843	2135	13	36	1.2 (0.7, 2.2)	2	17	0.4 (0.1, 1.6)
227. *Laboratory Products* *	849	2146	7	26	1.1 (0.5, 2.2)	5	17	1.3 (0.5, 3.3)
228. Fertilizers	797	2052	57	112	1.2 (0.9, 1.6)	24	54	0.9 (0.6, 1.4)
229. Pesticides	791	2039	54	116	1.0 (0.7, 1.3)	20	48	0.8 (0.5, 1.3)
230. Biocides	763	1933	78	211	0.8 (0.6, 1.0)	35	80	1.0 (0.7, 1.4)
231. Bleaches	849	2143	6	28	0.5 (0.2, 1.1)	3	10	0.7 (0.2, 2.3)

[a] Each row corresponds to a separate regression model to estimate the effects of that chemical alone, with adjustment only for the non-occupational confounders: age, income, ethnicity, cigarette use, alcohol use, respondent status, years of education, and recreational activity.
* An asterisk indicates that one or both of the estimates, which in this table use the cancer series as controls, differed statistically from the estimates using the electoral list series as controls (see appendices for estimates using the electoral list series).
Italics indicate the particular chemical will not appear in further tables of results for more complex regression models (see table 6-5 for explanation).

7.4 Comparison of estimates from different modeling strategies

One objective of the thesis was to compare the results of the semi-Bayes models to the results of the more conventional modeling approaches. Given the very large number of chemicals that were analyzed, it was necessary to summarize and synthesize the results for the purpose of inter-strategy comparisons. This section shows the distributions of the ensemble of beta estimates as well as cross-tabulated comparisons of results for a few informative pairings of models.

7.4.1 Distribution of estimates obtained from the modeling strategies

Table 7-4 provides the characteristics for the estimates of beta (log RR) for the ensemble of 184 parameters in the models for any level of exposure and of 146 parameters in the models of the substantial level of exposure. Although all 231 parameters were estimated in strategies 1, 2, and 3 (see Table 6-6), for comparison purposes, only the 184 or 146 parameters found in the other strategies were used for this table.

The location of the distribution of estimates (the mean and median of the regression coefficients) indicates that strategies 1 and 2 resulted in beta estimates that tended to be generally higher than estimates from the other strategies. The estimates from strategy 3, which adjusted for seven currently suspected lung carcinogens, were centered closer to zero. This shift partly suggests that strategies 1 and 2 likely overestimated several effects, and that this could be due to uncontrolled confounding from other occupational chemicals. Most of the distributions of estimates obtained from the higher numbered strategies were centered close to zero, which is what would be expected if many (or most) of the chemicals being assessed were not causally related to lung cancer.

Strategy 4 used a separate regression model for each chemical, with the data-dependent selection of other chemicals as confounders. For any level of exposure, the number of other chemicals included as confounders ranged from 46 to 57 with an average of 53. At the substantial level of exposure, the numbers ranged from 24 to 36 with an average of 31.

In modeling strategy 5, using an automatic forward selection strategy, the beta estimates and their standard errors were assumed to equal zero for those chemicals which did not satisfy the entry criterion (P-value ≤ 0.25). With this strategy, at any level of exposure,

55 of 184 chemicals remained in the model, and at the substantial level of exposure, 31 of 146 chemicals remained in the model. The low average standard error and very high kurtosis in Table 7-4 reflect the large number of zeroes that correspond to estimates for chemicals not entered in the model.

Table 7-4: Characteristics of the distributions of the ensemble of beta estimates for several approaches to modeling the occupational chemicals

Any level of exposure (K=184)	*Model strategy* [a]							
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Mean beta	0.14	0.11	0.01	0.01	0.01	-0.02	0.00	0.01
Median beta	0.20	0.12	0.02	0.05	0.00	0.00	0.01	0.01
Standard deviation of betas	0.29	0.30	0.30	0.44	0.37	0.53	0.25	0.24
Average of estimated standard errors	0.23	0.25	0.26	0.33	0.10	0.43	0.31	0.30
Skewness of betas	-0.83	-0.39	-0.51	-0.56	-0.71	-0.01	0.12	0.19
Kurtosis of betas	1.36	0.46	0.57	2.26	7.05	2.50	0.38	0.32
Lowest RR estimate [b]	0.36	0.44	0.39	0.14	0.14	0.13	0.52	0.50
Highest RR estimate	2.04	2.24	1.96	3.76	3.67	7.53	2.34	2.10
Substantial level of exposure (K=146)	*Model strategy* [a]							
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Mean beta	0.21	0.17	0.07	0.03	0.03	0.03	0.04	0.05
Median beta	0.20	0.16	0.05	-0.01	0.00	0.06	0.04	0.02
Standard deviation of betas	0.37	0.39	0.39	0.50	0.32	0.64	0.23	0.27
Average of estimated standard errors	0.36	0.39	0.40	0.45	0.10	0.61	0.38	0.38
Skewness of betas	0.14	0.54	0.61	-0.08	-0.07	-0.09	0.09	0.32
Kurtosis of betas	0.27	0.34	0.58	0.17	7.07	0.58	-0.65	-0.06
Lowest RR estimate [b]	0.45	0.49	0.44	0.25	0.25	0.18	0.63	0.55
Highest RR estimate	3.44	3.71	3.47	3.60	3.60	5.62	1.76	2.21

[a] Strategies 2-8 include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: (1) age-adjusted only, (2) adjustment only for non-occupational confounders, (3) adjustment for seven currently suspected lung carcinogens, (4) automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: (5) automatic forward selection of which chemicals to include using $P<0.25$, (6) all chemicals included, (7) semi-Bayes shrinkage of estimates from strategy 6 using a common prior, (8) semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] RR is rate ratio, the exponentiated beta.

As more parameters were added to the models, the standard errors of the betas predictably increased. However, the shrinkage from the semi-Bayes models meant that even with hundreds of parameters being estimated, strategies 7 and 8 resulted in standard errors that were on average similar to or lower than those from the simpler strategies.

To assess how similar was the ranking of the point estimates from different modeling strategies, Table 7-5 shows non-parametric rank correlation coefficients for the ensemble of estimates with each pairing of models. The Spearman correlation coefficient measures the correlation of the ranks of the estimates produced by the two models being compared. The results of Table 7-5 suggest that adding the non-occupational confounders to the model (strategy 2) and adding the seven suspected carcinogens to the model (strategy 3) both had some influence on the parameter estimates for the occupational chemicals, but it was the mutual adjustment for many occupational chemicals (strategies 4 through 8) that changed the ranking of the estimates appreciably. The estimates from strategy 7 remained highly correlated with the estimates of strategy 6, likely because the semi-Bayes shrinkage did not change the relative magnitude of too many of the estimates. Estimates from modeling strategy 5 tended to have the lowest correlation with estimates from other strategies. This undoubtedly was due to the point estimates for many chemicals being set to a value of zero.

Table 7-5: Spearman correlation coefficients among the logistic beta estimates from several modeling strategies

Any level of exposure (K=184)	*Modeling strategy* [a]						
	2	3	4	5	6	7	8
1	0.89	0.81	0.54	0.47	0.50	0.60	0.58
2	1.00	0.95	0.58	0.50	0.55	0.66	0.64
3		1.00	0.63	0.51	0.61	0.72	0.69
4			1.00	0.68	0.84	0.90	0.88
5				1.00	0.60	0.67	0.66
6					1.00	0.95	0.94
7						1.00	0.98
Substantial level of exposure (K=146)	*Modeling strategy* [a]						
	2	3	4	5	6	7	8
1	0.89	0.82	0.63	0.40	0.53	0.59	0.53
2	1.00	0.94	0.71	0.46	0.60	0.66	0.64
3		1.00	0.79	0.48	0.65	0.73	0.70
4			1.00	0.56	0.81	0.86	0.84
5				1.00	0.46	0.52	0.48
6					1.00	0.94	0.84
7						1.00	0.93

[a] Strategies 2-8 include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: (1) age-adjusted only, (2) adjustment only for non-occupational confounders, (3) adjustment for seven currently suspected lung carcinogens, (4) automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: (5) automatic forward selection of which chemicals to include using $P<0.25$, (6) all chemicals included, (7) semi-Bayes shrinkage of estimates from strategy 6 using a common prior, (8) semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability.

7.4.2 Further comparisons of the distributions of parameter estimates from different modeling approaches

An alternative representation of the behaviour of the different models would classify their estimates into broad categories of P-values, representing common views of evidence and statistical significance. Table 7-6 summarizes the distributions of the chemical-specific point estimates, using categories of direction of effect and of P-values. Each row displays the counts of estimates that fall on either side of unity, with further divisions by levels of P-values, using 0.05 to 0.15 to indicate a 'grey zone' of marginal statistical non-significance. Estimates with P-value ≥ 0.15 were considered definitely non-significant and were grouped together regardless of which side of unity they fell, because their departures from 1.0 were believed to reflect mostly sampling error.

In Table 7-6, strategies 1 and 2, and to a lesser extent, strategies 3 to 5, had far more elevated or statistically significant estimates than the strategies based on the single model approach that used all variables. Few elevated estimates were observed in the results of the semi-Bayes models. This was as expected. Shrinkage estimators should reduce the number of imprecise estimates that are likely spuriously large. Strategy 6, however, did not employ a Bayesian methodology, and alternative reasons for fewer elevated estimates are necessary. Some possible explanations might be that the simpler models did not account for the full extent of mutual confounding among the chemicals, and as such had inappropriately elevated estimates; or, alternatively, that the large model for strategy 6 involved 'over-adjustment' for the effects of several chemicals, for example if all the substances composed largely of silica had their estimates attenuated due to the inclusion of silica in the model. At any level of exposure, the addition of the seven suspected carcinogens in strategy 3 seemed to eliminate many of the elevated estimates from strategy 2, supporting the view of unaccounted confounding from, at the very least, the seven suspected carcinogens. That strategy 4 did not behave this way might be explained by the criterion used for confounder inclusion, namely a significant P-value, which would tend to only include chemicals associated with lung cancer irrespective of their roles as confounders. One noteworthy aspect of the results of strategy 4 is that while benzo(a)pyrene was occasionally selected for inclusion in the models, the rest of the seven suspected carcinogens of strategy 3 were not.

Table 7-6: Distributions of estimates from the different modeling strategies, according to direction of effect and P-value

Any level of exposure (K=184)		*RR<1.0 P < 0.05*	*RR<1.0 0.05 ≤P < 0.15*	*P ≥0.15*	*RR>1.0 0.05 ≤P < 0.15*	*RR>1.0 P < 0.05*
Modeling strategy [a]	*1*	0	7	108	24	45
	2	2	6	125	29	22
	3	5	7	153	14	5
	4	8	11	137	18	10
	5	8	8	149	8	11
	6	3	10	153	14	4
	7	1	4	172	4	3
	8	1	3	174	3	3
Substantial level of exposure (K=146)		*RR<1.0 P < 0.05*	*RR<1.0 0.05 ≤P < 0.15*	*P ≥0.15*	*RR>1.0 0.05 ≤P < 0.15*	*RR>1.0 P < 0.05*
Modeling strategy [a]	*1*	1	1	104	11	29
	2	0	1	111	16	18
	3	1	2	127	6	10
	4	4	6	117	12	7
	5	3	2	130	6	5
	6	2	4	123	14	3
	7	0	1	142	3	0
	8	0	1	139	6	0

[a] Strategies 2-8 include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: (1) age-adjusted only, (2) adjustment only for non-occupational confounders, (3) adjustment for seven currently suspected lung carcinogens, (4) automatic forward selection of other chemicals as confounders using P<0.25. All chemicals assessed in a single large regression model: (5) automatic forward selection of which chemicals to include using P<0.25, (6) all chemicals included, (7) semi-Bayes shrinkage of estimates from strategy 6 using a common prior, (8) semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability.

At the substantial level of exposure, only three estimates remained statistically significant at alpha 0.05 with the full model strategy 6. No estimates appeared statistically significant in the two semi-Bayes models (strategies 7 and 8), likely because the substantial-level exposure analysis resulted in estimates with relatively large variances, which the semi-Bayes shrinkage would have compensated for by shifting those estimates closer to the common mean.

A comparison was made of the observed numbers of statistically significant results with the number expect on the global null hypothesis. At the α=0.05 level, it can be expected that the number of statistically significant estimates would be 0.05*184=9.2 at the 'any exposure' level, and 0.05*146=7.3 at the 'substantial exposure' level. In Table 7-6, for any level of exposure, strategies 1 and 2 resulted in 45 and 24 statistically significant estimates, respectively. This was far more than expected by the simple calculation. While these departures from the two calculated values could be explained by a failure to account for any confounding in the former model, and confounding due to occupational carcinogens in the latter model, the results do provide some supporting evidence for the presence of true carcinogens in the study population, whether or not these were documented in the Montreal study. This is especially the case as the statistically significant estimates did not occur symmetrically above and below unity. The calculation for expected numbers, however, would be inappropriate if applied to the semi-Bayes results, because spuriously large or chance estimates would have already been corrected with the semi-Bayes estimator.

While Table 7-6 shows how the distribution of results compare across strategies, similar looking distributions do not necessarily mean the same chemicals were earmarked as statistically significant. The following four tables compare the distributions with the same tabulations of direction of effect and ranges of P-value as above, but now with cross-tabulated results for comparisons between selected pairs of modeling strategies. These four comparisons were chosen to highlight the influences of a few specific characteristics of the different strategies. The distributions in Table 7-7 show what influence the addition of the seven suspected lung carcinogens had on the set of estimates. The bottom row shows eight chemicals whose estimates were significant with P-value<0.05 in model 2, but which became non-significant (P-value>0.15) in model 3,

after additional adjustment for confounding from the suspected risk factors. In total, 17 chemicals earmarked as statistically significant by model 2 were 'eliminated' in model 3. One should notice, however, that a change in P-value may not necessarily imply an important change in the point estimate. An example of one of the eight chemicals is mild steel dust, whose point estimates and 90% confidence intervals under models 2 and 3, respectively, were 1.3 (1.1, 1.7) and 1.2 (0.9, 1.5), indicating that the 'loss of significance' was due to a combination of a slightly lower point estimate with nearly constant standard error. Thus, contrary to what might be gleaned from the cross-tabulation, these estimates were actually very similar, which evokes the difficulty of interpreting results using a P-value metric, combining both the magnitude of the estimate and its precision.

Table 7-7: Contrasting the distribution of results from strategies 2 and 3, ANY level of exposure: the effect of adding seven chemicals as confounders, chosen *a priori*

		Strategy 3 [a]				
		RR<1.0 P < 0.05	RR<1.0 0.05 ≤ P < 0.15	P ≥0.15	RR>1.0 0.05 ≤ P < 0.15	RR>1.0 P < 0.05
Strategy 2 [a]	RR<1.0 P < 0.05	2	0	0	0	0
	RR<1.0 0.05 ≤ P < 0.15	3	3	0	0	0
	P ≥0.15	0	4	121	0	0
	RR>1.0 0.05 ≤ P < 0.15	0	0	24	5	0
	RR>1.0 P < 0.05	0	0	8	9	5

[a] Each chemical in a separate regression model: Strategy 2 adjusts only for non-occupational confounders, and strategy 3 further adjusts for seven currently suspected lung carcinogens.

On the preventive side of the scale, and contrary to what would be expected of occupational carcinogens, three of the 'grey zone' estimates in model 2 became statistically significant with the addition of the seven suspected lung carcinogens in

strategy 3. On close examination, the estimates for these three chemicals were shifted downward to a degree that is greater than the increase in variance from adding variables to the model. For example, one of the chemicals is polyvinyl chlorate, whose point estimates and 90% confidence intervals under models 2 and 3, respectively, were 0.7 (0.4, 1.1) and 0.6 (0.3, 1.0).

Table 7-8 compares estimates that had been adjusted only for the seven currently suspected or recognized lung carcinogens, chosen *a priori* (strategy 3), with the estimates from the model where all chemicals were adjusted for each other, involving no pre-selection (strategy 6).

Table 7-8: Contrasting the distribution of results from strategies 3 and 6, ANY level of exposure: the effect of simultaneously adjusting for all chemicals

		Strategy 6 [a]				
		RR<1.0 P < 0.05	RR<1.0 0.05 ≤ P < 0.15	P ≥0.15	RR>1.0 0.05 ≤ P < 0.15	RR>1.0 P < 0.05
Strategy 3 [a]	RR<1.0 P < 0.05	1	1	3	0	0
	RR<1.0 0.05 ≤ P < 0.15	0	1	6	0	0
	P ≥0.15	2	8	131	9	3
	RR>1.0 0.05 ≤ P < 0.15	0	0	11	3	0
	RR>1.0 P < 0.05	0	0	2	2	1

[a] Strategy 3: Each chemical in a separate regression model, with adjustment for non-occupational confounders and seven currently suspected lung carcinogens. Strategy 6: All chemicals in a single model with further adjustment for non-occupational confounders.

The cross-tabulation shows that several estimates from the two models were quite different. The confidence intervals in strategy 6 were on average 65% wider than those of strategy 3, somewhat complicating the interpretation of the table. In principle, strategy 6

should have less confounding bias than strategy 3 due to the completeness of the mutual adjustment among the chemicals, but this would be offset by the far greater variance of these estimates and the possibility of over-adjustment of certain estimates.

Table 7-9 shows the results of taking the maximum likelihood estimates in a single regression model (strategy 6), and pulling them back toward close to unity with a semi-Bayes shrinkage estimator (strategy 7). There was a good amount of agreement between the two approaches. When the results did differ, many of the 'grey zone' estimates in strategy 6 became clearly statistically non-significant (P-value > 0.15) after accounting for their imprecision.

Table 7-9: Contrasting the distribution of results from strategies 6 and 7, ANY level of exposure: the effect of shrinkage to handle imprecision in strategy 6

		Strategy 7 [a]				
		RR<1.0 P < 0.05	RR<1.0 0.05 ≤ P < 0.15	P ≥0.15	RR>1.0 0.05 ≤ P < 0.15	RR>1.0 P < 0.05
Strategy 6 [a]	RR<1.0 P < 0.05	1	2	0	0	0
	RR<1.0 0.05 ≤ P < 0.15	0	2	8	0	0
	P ≥0.15	0	0	153	0	0
	RR>1.0 0.05 ≤ P < 0.15	0	0	11	3	0
	RR>1.0 P < 0.05	0	0	0	1	3

[a] Strategy 6: All chemicals in a single regression model. Strategy 7: Estimates from model 6 shrunk to a common prior, under assumption of all parameters exchangeable.

Finally, Table 7-10 shows how the semi-Bayes models were affected by replacing the common prior, where all chemical effects were assumed exchangeable (strategy 7), with a prior that separated chemicals into sets of exchangeable effects based on chemical and

physical properties and previous evidence of lung carcinogenicity (strategy 8). There was close agreement between the point estimates and P-values yielded by these two approaches. Most estimates changed very little. One of the few chemicals influenced by the inclusion of categories of exchangeability was lead chromate, whose point estimates and 90% interval estimates under models 7 and 8, respectively, were 1.1 (0.6, 2.1) and 1.4 (0.7, 2.8), though inference on its effect would hardly be altered.

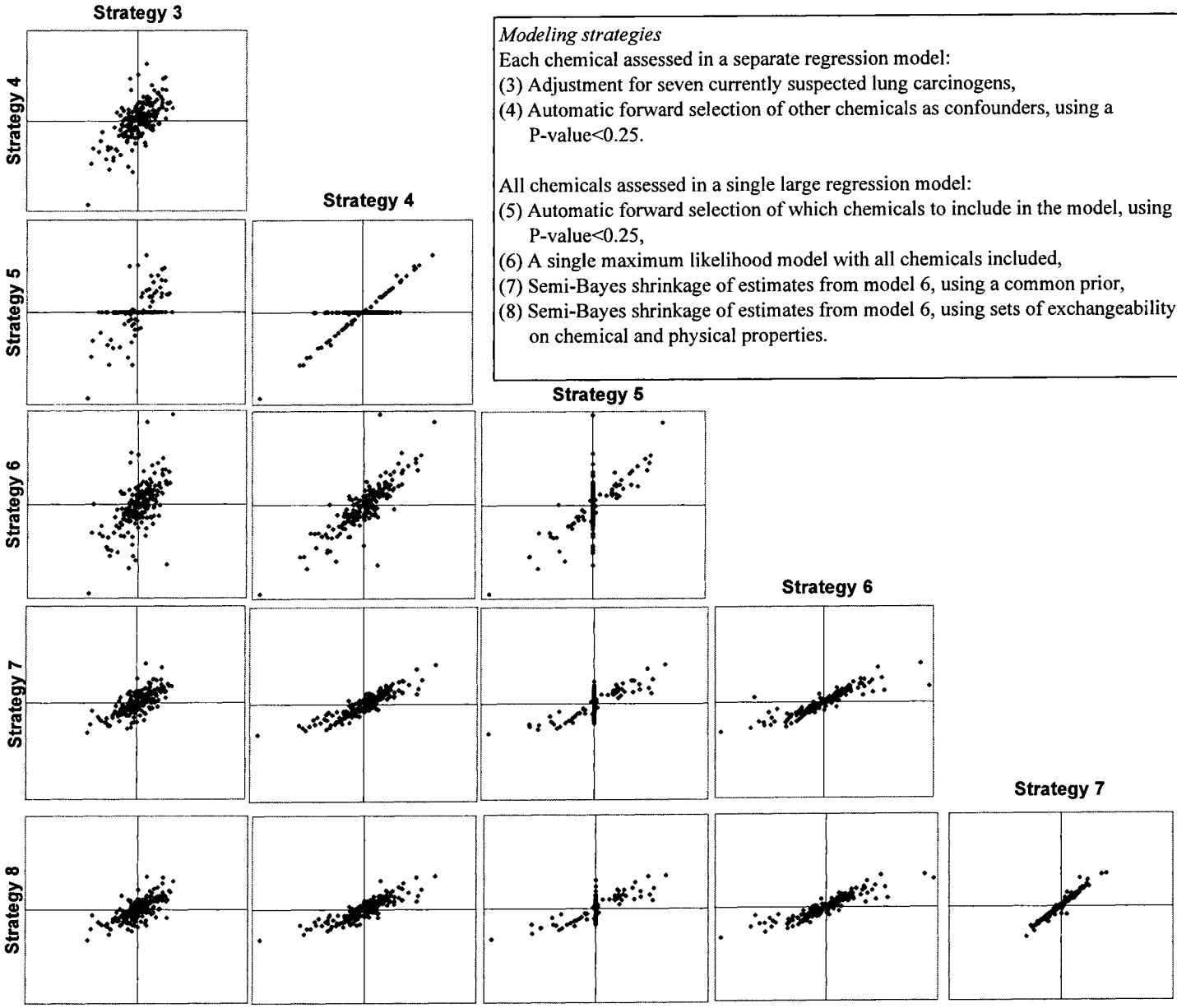
Table 7-10: Contrasting the distribution of results from strategies 7 and 8, ANY level of exposure: the effect of specifying categories of exchangeability based on expert opinion

		Strategy 8 [a]				
		RR<1.0 P < 0.05	RR<1.0 0.05 ≤ P < 0.15	P ≥0.15	RR>1.0 0.05 ≤ P < 0.15	RR>1.0 P < 0.05
Strategy 7 [a]	RR<1.0 P < 0.05	1	0	0	0	0
	RR<1.0 0.05 ≤ P < 0.15	0	2	2	0	0
	P ≥0.15	0	1	171	0	0
	RR>1.0 0.05 ≤ P < 0.15	0	0	1	3	0
	RR>1.0 P < 0.05	0	0	0	0	3

[a] Strategy 7: Estimates from model 6 shrunk to a common prior, under assumption of all parameters exchangeable. Strategy 8: Estimates from model 6 shrunk within sets of exchangeability, based on shared chemical and physical properties.

For a last comparison of the distributions of point estimates from the various modeling strategies, Figure 7-1 displays scatter plots of the logistic beta estimates for a selection of models, and is a graphical representation of much of the results shown above. Strategies 1 and 2 were dropped at this point, partly for space considerations but also because the other strategies had preferable properties, and these were rather simplistic approaches.

Figure 7-1: Scatter plots of logistic beta estimates for different pairings of modeling strategies, ANY level of exposure



* All axes represent a range of -2.1 to +2.1 on the scale for logistic coefficients

Neither these plots nor most of the results of this section suggest which of the models are better than others, but the results do show varying levels of discordance among the empirical estimates. The closer the spread of points lies to the bottom-left, upper-right diagonal of the plots in Figure 7-1, the more similar these models behaved in terms of their point estimation.

The estimates of strategy 3 tended to be quite different from the estimates of the more complex strategies, and, in addition, the point estimates from strategy 3 showed only moderate agreement with estimates from the semi-Bayes analysis (strategy 8). In contrast, strategies 4 and 6 resulted in estimates that were proportional to, but often more extreme than, the estimates from strategies 7 and 8.

Strategy 5 produced distinct plots because many of the estimates were set to zero as a result of the variable being deleted (P-value>0.25) from the regression model. The plot comparing strategies 4 and 5 shows the estimates falling nearly perfectly along two lines. For the estimates along the diagonal, the automatic selection procedures produced regression equations that were very similar. In other words, the same chemicals tended to be retained as confounders in each of the many regression equations of strategy 4, and these were nearly the same chemicals that were selected for inclusion in the single regression equation of strategy 5.

Strategies 7 and 8 produced very similar plots. Adding the categories of exchangeability to the semi-Bayes models had only minor influence on the estimates.

7.5 Models accounting for potential mutual confounding among the chemicals

Accounting for potential mutual confounding among the occupational chemicals was a primary objective of the thesis. Strategies were considered that included adjustment for the effects of only a few other chemicals, while other strategies involved the simultaneous adjustment for the effects of a large number of chemicals. Table 7-11 presents results for the 184 chemicals found in the larger modeling strategies. Only four of the eight strategies were considered here, at two levels of exposure. These strategies represent differing degrees of control for any mutual confounding among the chemicals (see section 6.9). Moving across the columns, the models begin by choosing seven currently suspected lung carcinogens as *a priori* confounders (strategy 3), choosing many confounders via an automatic P-value selection (strategy 4), and avoiding pre-selection by including all the chemicals simultaneously (strategies 7 and 8). As for the latter, it would be natural to include in this table the estimates from strategy 6, the unrestricted logistic regression, but those estimates were imprecise and not intended for interpretation. Instead, the semi-Bayes estimates are presented. These models used the estimates from the model of strategy 6 and shrunk them using a second-level model.

In Table 7-11, the substances are grouped by the categories of exchangeability used in semi-Bayes strategy 8. This allowed some visual assessment of whether sets of chemicals had similar effects or not. So, for example, all chemicals that are polypeptides were grouped together in the same sub-section of the table. A few substances were repeated several times if they belonged to several categories, though the row of values reported will be identical. For example, the same estimates for asbestos (5) are found under the categories previous evidence, fibrous inorganic dusts, and silica-containing compounds. Following each category label is a descriptor of which levels of exposure that category was used in. For example, the chromates category has only the word 'any' following the label. This means that the covariate for the chromates category was not used in the second-level regression model at the substantial level of exposure, because too few chemicals remained in that set.

Table 7-11 is designed so that by "reading downward" in the columns of semi-Bayes results (strategies 7 and 8), it is possible to know which chemicals were adjusted for each other in the single logistic model. For example, lead chromate (40) was included in the model for any level of exposure, but not in the model for substantial-level exposure effects, due to sparse data at that level.

The last section of the table, titled "not in a category", lists the estimates for chemicals which were included in the semi-Bayes regression models but not in any particular category of exchangeability. These estimates were still shrunk toward the overall prior mean, close to unity, but they tended not to share chemical and physical properties with other chemicals on the list.

Previous evidence (of lung carcinogenicity) was not represented with a dichotomous covariate in the models, as the other categories were. For the purpose of this table, however, the chemicals listed in that section are those that had non-zero values for previous evidence.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
Previous evidence [Any, Substantial] [b]								
5. Asbestos	1.0 (0.8,1.3)	1.5 (1.0,2.4)	1.1 (0.8,1.4)	2.0 (1.2,3.3)	1.0 (0.7, 1.3)	1.4 (0.9, 2.3)	1.0 (0.7,1.3)	1.4 (0.9,2.4)
6. Crystalline Silica	1.2 (1.0,1.5)	1.6 (1.2,2.2)	0.9 (0.7,1.2)	1.6 (1.1,2.2)	1.0 (0.8, 1.4)	1.3 (0.8, 1.9)	1.0 (0.8,1.3)	1.3 (0.9,1.9)
34. Iron Oxides	0.9 (0.7,1.1)	0.7 (0.5,1.1)	0.8 (0.6,1.1)	0.7 (0.4,1.1)	0.8 (0.5, 1.1)	0.8 (0.5, 1.3)	0.7 (0.5,1.0)	0.8 (0.5,1.2)
40. Lead Chromate	0.7 (0.4,1.2)	1.1 (0.2,5.6)	1.3 (0.7,2.6)		1.1 (0.6, 2.0)		1.4 (0.7,2.8)	
76. Soot	1.0 (0.7,1.4)	1.6 (0.9,2.7)	1.2 (0.8,1.8)	1.1 (0.6,1.9)	0.9 (0.6, 1.4)	0.9 (0.5, 1.7)	0.9 (0.6,1.4)	1.0 (0.5,1.8)
97. Coal Gas	0.6 (0.3,1.1)	1.4 (0.3,6.8)	0.5 (0.2,1.2)		0.7 (0.3, 1.5)		0.7 (0.4,1.4)	
104. Chromium Fumes	2.0 (1.2,3.2)	2.1 (0.7,6.6)	1.3 (0.8,2.2)	1.7 (0.6,5.1)	1.4 (0.7, 2.9)	1.4 (0.7, 3.1)	1.9 (0.6,6.1)	1.6 (0.7,4.0)
107. Nickel Fumes	1.7 (1.1,2.8)		1.3 (0.8,2.1)		1.1 (0.5, 2.3)		0.8 (0.3,2.5)	
117. Diesel Eng.Emissions	1.1 (0.9,1.4)	1.1 (0.9,1.5)	1.1 (0.9,1.4)	1.4 (1.0,1.9)	1.1 (0.9, 1.4)	1.2 (0.8, 1.7)	1.1 (0.9,1.4)	1.2 (0.8,1.7)
132. Sulphuric Acid	0.9 (0.7,1.2)	0.7 (0.4,1.3)	0.7 (0.5,1.1)	0.7 (0.4,1.3)	0.8 (0.5, 1.2)	0.8 (0.4, 1.6)	0.8 (0.5,1.2)	0.8 (0.4,1.5)
163. Coal Tar and Pitch	0.9 (0.5,1.4)	1.0 (0.5,2.3)	0.8 (0.5,1.4)	1.0 (0.5,2.1)	0.9 (0.5, 1.6)	1.1 (0.5, 2.1)	0.9 (0.6,1.5)	1.1 (0.6,2.0)
181. Beryllium Compounds	0.7 (0.2,2.1)		0.5 (0.1,1.9)		0.8 (0.4, 1.8)		0.8 (0.4,1.7)	
193. Arsenic Compounds	0.7 (0.5,1.1)	0.9 (0.5,1.7)	0.9 (0.6,1.5)	0.7 (0.4,1.5)	0.8 (0.5, 1.3)	0.9 (0.5, 1.8)	0.7 (0.4,1.3)	1.2 (0.5,2.7)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using P<0.25. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
195. Cadmium Compounds	1.5 (0.7,3.0)	1.4 (0.5,4.4)	0.9 (0.4,2.2)	0.9 (0.2,3.2)	0.9 (0.5, 1.8)	1.1 (0.5, 2.4)	1.0 (0.5,2.2)	1.1 (0.4,2.8)
214. PAH (Any)	1.0 (0.9,1.3)	1.2 (0.8,1.8)	0.9 (0.8,1.1)	1.1 (0.7,1.5)	1.0 (0.7, 1.3)	1.3 (0.8, 2.0)	1.0 (0.7,1.3)	1.2 (0.8,1.9)
219. Benzo(a)pyrene	1.1 (0.9,1.4)	1.0 (0.6,1.6)	1.0 (0.8,1.3)	1.2 (0.7,2.2)	1.1 (0.8, 1.5)	0.9 (0.5, 1.7)	1.1 (0.8,1.5)	1.0 (0.5,1.7)
Polypeptides [Any, Substantial] [b]								
43. Wool Fibres	0.9 (0.7,1.3)	0.8 (0.5,1.2)	1.0 (0.7,1.5)	0.9 (0.5,1.3)	1.0 (0.6, 1.5)	0.9 (0.5, 1.6)	1.0 (0.6,1.5)	1.0 (0.6,1.8)
47. Fur Dust	1.3 (0.7,2.3)	0.9 (0.4,2.1)	1.4 (0.8,2.7)	1.0 (0.4,2.3)	1.2 (0.7, 2.1)	1.1 (0.6, 2.2)	1.2 (0.7,2.1)	1.2 (0.6,2.4)
48. Hair Dust	1.0 (0.5,2.0)	1.1 (0.5,2.9)	0.9 (0.4,1.9)	1.2 (0.5,3.1)	1.0 (0.5, 2.0)	1.3 (0.6, 2.7)	1.0 (0.5,2.0)	1.4 (0.6,3.1)
51. Leather Dust	0.8 (0.5,1.2)	0.9 (0.4,1.7)	1.0 (0.6,1.7)	1.1 (0.5,2.3)	0.9 (0.6, 1.6)	1.2 (0.6, 2.4)	1.0 (0.6,1.6)	1.3 (0.7,2.6)
Polysaccharides [Any, Substantial] [b]								
42. Cotton Dust	0.9 (0.7,1.2)	0.9 (0.6,1.3)	1.1 (0.8,1.4)	0.9 (0.6,1.3)	1.0 (0.7, 1.5)	0.9 (0.6, 1.5)	1.0 (0.7,1.5)	0.9 (0.6,1.5)
44. Wood Dust	1.1 (0.9,1.4)	1.3 (1.0,1.6)	1.1 (0.9,1.4)	1.2 (0.9,1.6)	1.1 (0.9, 1.4)	1.1 (0.8, 1.5)	1.1 (0.9,1.4)	1.1 (0.8,1.5)
45. Grain Dust	0.9 (0.7,1.2)	0.9 (0.6,1.4)	0.8 (0.6,1.2)	1.1 (0.7,1.8)	0.9 (0.6, 1.3)	1.1 (0.7, 1.7)	0.9 (0.6,1.3)	1.0 (0.7,1.6)
46. Flour Dust	1.0 (0.7,1.4)	0.7 (0.4,1.2)	1.1 (0.7,1.6)	0.7 (0.4,1.2)	1.0 (0.7, 1.5)	0.8 (0.5, 1.3)	1.0 (0.7,1.5)	0.8 (0.5,1.3)
49. Starch Dust	1.5 (0.8,2.7)	1.4 (0.5,3.6)	1.3 (0.6,2.5)	0.8 (0.3,2.4)	1.2 (0.7, 2.1)	1.0 (0.5, 2.2)	1.2 (0.7,2.1)	1.0 (0.5,2.0)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
50. Sugar Dust	1.6 (0.9,3.0)	1.3 (0.5,3.4)	1.6 (0.8,3.3)	1.6 (0.6,4.8)	1.5 (0.8, 2.7)	1.2 (0.6, 2.5)	1.5 (0.8,2.6)	1.1 (0.6,2.3)
52. Tobacco Dust	1.1 (0.5,2.4)	0.8 (0.2,3.0)	1.2 (0.5,2.6)	0.9 (0.2,3.2)	1.1 (0.6, 2.0)	0.9 (0.4, 2.1)	1.1 (0.6,2.1)	0.9 (0.4,1.9)
75. Cellulose	1.1 (0.8,1.5)	1.0 (0.6,1.8)	1.4 (1.0,1.9)	0.9 (0.5,1.6)	1.2 (0.9, 1.7)	1.0 (0.6, 1.6)	1.2 (0.9,1.7)	0.9 (0.5,1.6)
Fibrous inorganic dusts [Any, Substantial] [b]								
2. Inorg.Insul.Dust	1.0 (0.7,1.3)	1.0 (0.6,1.5)	0.8 (0.6,1.2)	0.6 (0.3,1.2)	0.8 (0.5, 1.2)	0.7 (0.4, 1.2)	0.8 (0.5,1.2)	0.7 (0.4,1.2)
5. Asbestos	1.0 (0.8,1.3)	1.5 (1.0,2.4)	1.1 (0.8,1.4)	2.0 (1.2,3.3)	1.0 (0.7, 1.3)	1.4 (0.9, 2.3)	1.0 (0.7,1.3)	1.4 (0.9,2.4)
9. Glass Fibres	0.8 (0.6,1.1)	0.7 (0.3,1.3)	0.9 (0.5,1.6)	0.7 (0.3,1.7)	0.8 (0.5, 1.3)	0.9 (0.4, 1.7)	0.8 (0.5,1.3)	0.8 (0.4,1.6)
19. Mineral Wool Fibres	1.0 (0.7,1.3)	0.6 (0.3,1.2)	1.1 (0.7,1.9)	0.6 (0.3,1.5)	1.2 (0.7, 1.9)	0.8 (0.4, 1.6)	1.2 (0.7,1.9)	0.8 (0.4,1.5)
Silica containing compounds [Any, Substantial] [b]								
3. Excavation Dust	1.3 (1.0,1.8)	1.5 (1.1,2.2)	1.3 (1.0,1.7)	1.2 (0.8,1.8)	1.2 (0.9, 1.7)	1.3 (0.9, 2.0)	1.2 (0.9,1.7)	1.3 (0.9,2.0)
5. Asbestos	1.0 (0.8,1.3)	1.5 (1.0,2.4)	1.1 (0.8,1.4)	2.0 (1.2,3.3)	1.0 (0.7, 1.3)	1.4 (0.9, 2.3)	1.0 (0.7,1.3)	1.4 (0.9,2.4)
6. Crystalline Silica	1.2 (1.0,1.5)	1.6 (1.2,2.2)	0.9 (0.7,1.2)	1.6 (1.1,2.2)	1.0 (0.8, 1.4)	1.3 (0.8, 1.9)	1.0 (0.8,1.3)	1.3 (0.9,1.9)
7. Portland Cement	1.3 (0.9,1.7)	1.2 (0.8,1.7)	1.4 (1.0,2.0)	1.1 (0.7,1.7)	1.3 (0.9, 1.9)	1.2 (0.8, 1.9)	1.3 (0.9,1.9)	1.2 (0.8,1.8)
8. Glass Dust	1.9 (1.0,3.3)	0.7 (0.3,1.8)	1.5 (0.8,2.9)	0.6 (0.2,1.7)	1.4 (0.8, 2.5)	0.9 (0.4, 1.8)	1.5 (0.9,2.6)	1.0 (0.5,1.9)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
9. Glass Fibres	0.8 (0.6,1.1)	0.7 (0.3,1.3)	0.9 (0.5,1.6)	0.7 (0.3,1.7)	0.8 (0.5, 1.3)	0.9 (0.4, 1.7)	0.8 (0.5,1.3)	0.8 (0.4,1.6)
10. Industrial Talc	0.7 (0.5,1.1)	0.6 (0.3,1.3)	0.9 (0.5,1.5)	0.6 (0.2,1.3)	0.8 (0.5, 1.3)	0.8 (0.4, 1.7)	0.8 (0.5,1.4)	0.9 (0.5,1.8)
11. Brick Dust	0.8 (0.6,1.2)	0.7 (0.3,1.5)	1.2 (0.7,2.0)	0.7 (0.3,1.5)	1.2 (0.7, 2.0)	0.8 (0.4, 1.5)	1.2 (0.8,2.0)	0.8 (0.4,1.6)
12. Clay Dust	1.6 (1.0,2.6)	1.1 (0.5,2.4)	2.1 (1.2,3.6)	1.2 (0.5,3.0)	1.6 (1.0, 2.8)	1.2 (0.6, 2.4)	1.8 (1.0,3.2)	1.2 (0.6,2.7)
13. Concrete Dust	1.0 (0.8,1.4)	1.3 (0.9,1.9)	1.0 (0.7,1.5)	1.5 (1.0,2.3)	1.0 (0.7, 1.4)	1.3 (0.8, 1.9)	1.0 (0.7,1.5)	1.3 (0.8,1.9)
19. Mineral Wool Fibres	1.0 (0.7,1.3)	0.6 (0.3,1.2)	1.1 (0.7,1.9)	0.6 (0.3,1.5)	1.2 (0.7, 1.9)	0.8 (0.4, 1.6)	1.2 (0.7,1.9)	0.8 (0.4,1.5)
23. Cosmetic Talc	1.5 (0.8,2.7)	0.4 (0.1,1.6)	1.9 (1.0,3.7)		1.6 (0.9, 3.0)		1.6 (0.8,3.2)	
27. Silicon Carbide	0.9 (0.7,1.3)	1.0 (0.5,2.1)	0.7 (0.4,1.1)	0.8 (0.3,2.0)	0.7 (0.5, 1.1)	1.2 (0.6, 2.3)	0.8 (0.5,1.2)	1.1 (0.6,2.2)
Metal dusts (excluding oxides) [Any, Substantial] [b]								
14. Bronze Dust	0.8 (0.4,1.6)	0.7 (0.3,1.9)	0.6 (0.3,1.4)	0.5 (0.2,1.4)	0.8 (0.4, 1.4)	0.8 (0.4, 1.7)	0.7 (0.4,1.5)	0.6 (0.2,1.8)
15. Brass Dust	1.4 (0.9,2.3)	1.5 (0.7,3.1)	1.2 (0.7,2.1)	1.3 (0.6,3.0)	1.2 (0.7, 2.1)	1.3 (0.7, 2.6)	1.3 (0.8,2.3)	2.0 (0.9,4.5)
16. Stainless Steel Dust	1.4 (1.0,1.9)	1.5 (0.9,2.6)	1.2 (0.8,1.9)	1.5 (0.8,2.7)	1.1 (0.7, 1.6)	1.3 (0.7, 2.4)	1.1 (0.7,1.7)	1.4 (0.7,2.7)
17. Mild Steel Dust	1.2 (1.0,1.5)	1.1 (0.8,1.5)	1.2 (0.9,1.6)	1.2 (0.8,1.7)	1.2 (0.9, 1.6)	1.1 (0.7, 1.6)	1.2 (0.9,1.6)	1.2 (0.8,1.8)
21. Aluminium Alloy Dust	1.4 (1.0,1.9)	1.1 (0.7,1.8)	1.5 (1.0,2.3)	1.6 (0.9,2.6)	1.3 (0.9, 1.9)	1.2 (0.7, 2.1)	1.3 (0.9,1.9)	1.2 (0.7,2.2)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
33. Iron Dust	0.9 (0.6,1.4)	0.6 (0.3,1.2)	0.9 (0.5,1.4)	0.5 (0.2,1.0)	0.9 (0.6, 1.4)	0.7 (0.4, 1.4)	0.8 (0.5,1.3)	0.6 (0.3,1.1)
35. Copper Dust	1.2 (0.8,1.7)	2.6 (1.3,5.3)	1.2 (0.7,1.9)	2.1 (1.0,4.5)	1.2 (0.7, 1.9)	1.7 (0.9, 3.2)	1.2 (0.8,2.0)	1.9 (0.9,3.9)
36. Zinc Dust	1.4 (0.9,2.3)	3.5 (1.1,11.1)	1.6 (0.9,2.9)	2.7 (0.7,10.7)	1.2 (0.7, 2.1)	1.3 (0.6, 3.1)	1.2 (0.7,2.1)	1.7 (0.6,4.9)
Metal oxide dusts [Any, Substantial] [b]								
26. Alumina	1.2 (0.9,1.5)	1.4 (0.9,2.2)	1.6 (1.2,2.3)	1.7 (0.9,3.0)	1.4 (1.0, 1.9)	1.6 (0.9, 2.8)	1.4 (1.0,1.9)	1.7 (1.0,3.1)
29. Calcium Oxide	1.0 (0.7,1.4)	1.1 (0.7,1.9)	0.7 (0.5,1.1)	1.0 (0.6,1.8)	0.8 (0.6, 1.2)	0.9 (0.5, 1.6)	0.9 (0.6,1.2)	1.1 (0.6,1.9)
32. Titanium Dioxide	0.9 (0.6,1.3)	1.8 (0.6,5.0)	1.1 (0.6,2.1)	1.5 (0.5,4.8)	1.2 (0.7, 2.0)	1.3 (0.6, 2.8)	1.1 (0.6,2.0)	1.6 (0.7,3.7)
34. Iron Oxides	0.9 (0.7,1.1)	0.7 (0.5,1.1)	0.8 (0.6,1.1)	0.7 (0.4,1.1)	0.8 (0.5, 1.1)	0.8 (0.5, 1.3)	0.7 (0.5,1.0)	0.8 (0.5,1.2)
37. Zinc Oxide	1.0 (0.7,1.6)	0.8 (0.2,2.8)	1.5 (0.9,2.7)		1.3 (0.7, 2.2)		1.3 (0.7,2.3)	
38. Lead Oxides	1.5 (0.9,2.5)	1.9 (0.8,4.9)	1.3 (0.7,2.5)	1.7 (0.7,4.7)	1.3 (0.8, 2.3)	1.4 (0.7, 3.0)	1.3 (0.7,2.4)	1.6 (0.6,3.8)
Metal oxide fumes [Any, Substantial] [b]								
98. Gas Welding Fumes	1.3 (1.0,1.7)	1.4 (1.0,2.0)	1.1 (0.8,1.5)	1.0 (0.6,1.7)	1.3 (0.9, 1.8)	0.9 (0.5, 1.5)	1.2 (0.9,1.8)	0.9 (0.6,1.6)
99. Arc Welding Fumes	1.0 (0.8,1.3)	1.3 (0.9,1.9)	0.7 (0.5,0.9)	0.5 (0.3,0.9)	0.7 (0.4, 1.0)	0.7 (0.4, 1.3)	0.7 (0.5,1.0)	0.7 (0.4,1.2)
100. Soldering Fumes	1.0 (0.7,1.4)	1.2 (0.8,1.9)	1.0 (0.7,1.5)	1.0 (0.6,1.6)	0.9 (0.6, 1.3)	1.0 (0.6, 1.7)	0.9 (0.6,1.3)	1.0 (0.6,1.7)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using P<0.25. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
101. Metal Oxide Fumes	1.2 (1.0,1.4)	1.4 (1.1,1.9)	1.1 (0.8,1.5)	0.9 (0.6,1.3)	1.1 (0.7, 1.5)	1.3 (0.8, 2.1)	1.1 (0.7,1.5)	1.3 (0.8,2.1)
102. Aluminium Fumes	1.2 (0.7,1.9)	0.8 (0.4,1.9)	0.6 (0.3,1.2)	0.3 (0.1,0.7)	0.7 (0.4, 1.2)	0.7 (0.3, 1.4)	0.7 (0.4,1.3)	0.7 (0.3,1.6)
103. Calcium Oxide Fumes	1.1 (0.8,1.5)	1.2 (0.7,1.9)	0.7 (0.5,1.2)	0.6 (0.3,1.3)	0.8 (0.5, 1.5)	0.6 (0.3, 1.2)	0.9 (0.5,1.5)	0.7 (0.4,1.3)
104. Chromium Fumes	2.0 (1.2,3.2)	2.1 (0.7,6.6)	1.3 (0.8,2.2)	1.7 (0.6,5.1)	1.4 (0.7, 2.9)	1.4 (0.7, 3.1)	1.9 (0.6,6.1)	1.6 (0.7,4.0)
105. Manganese Fumes	1.4 (1.0,1.9)	3.3 (1.5,7.0)	1.0 (0.6,1.6)	1.8 (0.7,4.6)	1.1 (0.6, 1.8)	1.4 (0.7, 2.9)	1.2 (0.7,2.2)	2.2 (0.9,5.5)
106. Iron Fumes	1.2 (0.9,1.6)	1.4 (1.0,2.1)	0.9 (0.6,1.3)	0.8 (0.5,1.5)	0.8 (0.5, 1.3)	1.0 (0.5, 1.9)	0.7 (0.4,1.2)	0.6 (0.3,1.3)
107. Nickel Fumes	1.7 (1.1,2.8)		1.3 (0.8,2.1)		1.1 (0.5, 2.3)		0.8 (0.3,2.5)	
108. Copper Fumes	1.8 (1.2,2.7)	2.1 (1.1,3.9)	1.5 (0.9,2.4)	2.2 (1.0,4.5)	1.5 (0.9, 2.6)	1.3 (0.7, 2.6)	1.5 (0.9,2.7)	1.6 (0.7,3.6)
109. Zinc Fumes	1.3 (0.9,1.9)	2.6 (1.3,5.1)	0.9 (0.5,1.6)	1.5 (0.6,3.7)	0.9 (0.5, 1.5)	1.2 (0.6, 2.4)	0.9 (0.5,1.5)	1.4 (0.6,3.3)
110. Silver Fumes	1.2 (0.7,2.1)	1.4 (0.6,3.3)	0.6 (0.2,1.3)	1.1 (0.4,3.0)	0.8 (0.4, 1.5)	1.1 (0.5, 2.3)	0.8 (0.4,1.5)	1.0 (0.5,2.0)
111. Tin Fumes	1.4 (1.0,2.0)	1.0 (0.5,2.1)	1.1 (0.6,2.1)	0.5 (0.2,1.3)	1.0 (0.6, 1.8)	1.0 (0.5, 2.0)	0.8 (0.4,1.6)	0.6 (0.2,1.7)
112. Lead Fumes	1.3 (0.9,1.8)	0.8 (0.3,1.6)	1.0 (0.5,2.0)	0.5 (0.2,1.1)	1.1 (0.6, 1.9)	0.8 (0.4, 1.8)	1.2 (0.6,2.1)	0.9 (0.4,2.1)
130. Nitric Acid	0.8 (0.4,1.5)	0.5 (0.2,1.5)	0.7 (0.3,1.7)	0.5 (0.1,1.8)	0.8 (0.4, 1.5)	0.9 (0.4, 2.0)	0.8 (0.4,1.5)	0.8 (0.3,1.7)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using P<0.25. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
Heavy metal compounds [Any, Substantial] [b]								
14. Bronze Dust	0.8 (0.4,1.6)	0.7 (0.3,1.9)	0.6 (0.3,1.4)	0.5 (0.2,1.4)	0.8 (0.4, 1.4)	0.8 (0.4, 1.7)	0.7 (0.4,1.5)	0.6 (0.2,1.8)
15. Brass Dust	1.4 (0.9,2.3)	1.5 (0.7,3.1)	1.2 (0.7,2.1)	1.3 (0.6,3.0)	1.2 (0.7, 2.1)	1.3 (0.7, 2.6)	1.3 (0.8,2.3)	2.0 (0.9,4.5)
16. Stainless Steel Dust	1.4 (1.0,1.9)	1.5 (0.9,2.6)	1.2 (0.8,1.9)	1.5 (0.8,2.7)	1.1 (0.7, 1.6)	1.3 (0.7, 2.4)	1.1 (0.7,1.7)	1.4 (0.7,2.7)
35. Copper Dust	1.2 (0.8,1.7)	2.6 (1.3,5.3)	1.2 (0.7,1.9)	2.1 (1.0,4.5)	1.2 (0.7, 1.9)	1.7 (0.9, 3.2)	1.2 (0.8,2.0)	1.9 (0.9,3.9)
35. Copper Dust	1.2 (0.8,1.7)	2.6 (1.3,5.3)	1.2 (0.7,1.9)	2.1 (1.0,4.5)	1.2 (0.7, 1.9)	1.7 (0.9, 3.2)	1.2 (0.8,2.0)	1.9 (0.9,3.9)
38. Lead Oxides	1.5 (0.9,2.5)	1.9 (0.8,4.9)	1.3 (0.7,2.5)	1.7 (0.7,4.7)	1.3 (0.8, 2.3)	1.4 (0.7, 3.0)	1.3 (0.7,2.4)	1.6 (0.6,3.8)
39. Basic Lead Carb.	1.2 (0.7,1.9)	1.4 (0.3,6.1)	1.7 (1.0,3.0)		1.2 (0.6, 2.2)		1.2 (0.6,2.2)	
40. Lead Chromate	0.7 (0.4,1.2)	1.1 (0.2,5.6)	1.3 (0.7,2.6)		1.1 (0.6, 2.0)		1.4 (0.7,2.8)	
104. Chromium Fumes	2.0 (1.2,3.2)	2.1 (0.7,6.6)	1.3 (0.8,2.2)	1.7 (0.6,5.1)	1.4 (0.7, 2.9)	1.4 (0.7, 3.1)	1.9 (0.6,6.1)	1.6 (0.7,4.0)
107. Nickel Fumes	1.7 (1.1,2.8)		1.3 (0.8,2.1)		1.1 (0.5, 2.3)		0.8 (0.3,2.5)	
108. Copper Fumes	1.8 (1.2,2.7)	2.1 (1.1,3.9)	1.5 (0.9,2.4)	2.2 (1.0,4.5)	1.5 (0.9, 2.6)	1.3 (0.7, 2.6)	1.5 (0.9,2.7)	1.6 (0.7,3.6)
112. Lead Fumes	1.3 (0.9,1.8)	0.8 (0.3,1.6)	1.0 (0.5,2.0)	0.5 (0.2,1.1)	1.1 (0.6, 1.9)	0.8 (0.4, 1.8)	1.2 (0.6,2.1)	0.9 (0.4,2.1)
155. Leaded Gasoline	1.1 (0.8,1.3)	0.9 (0.6,1.3)	1.1 (0.9,1.5)	0.8 (0.6,1.2)	1.0 (0.7, 1.4)	0.9 (0.5, 1.4)	1.0 (0.7,1.5)	0.9 (0.6,1.6)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
193. Arsenic Compounds	0.7 (0.5,1.1)	0.9 (0.5,1.7)	0.9 (0.6,1.5)	0.7 (0.4,1.5)	0.8 (0.5, 1.3)	0.9 (0.5, 1.8)	0.7 (0.4,1.3)	1.2 (0.5,2.7)
195. Cadmium Compounds	1.5 (0.7,3.0)	1.4 (0.5,4.4)	0.9 (0.4,2.2)	0.9 (0.2,3.2)	0.9 (0.5, 1.8)	1.1 (0.5, 2.4)	1.0 (0.5,2.2)	1.1 (0.4,2.8)
Monocyclic aromatic hydrocarbons [Any, Substantial] [b]								
144. Benzene	1.0 (0.8,1.2)	0.8 (0.5,1.1)	1.3 (1.0,1.8)	0.8 (0.5,1.3)	1.2 (0.9, 1.8)	1.1 (0.6, 1.9)	1.2 (0.8,1.7)	1.0 (0.6,1.7)
145. Toluene	0.8 (0.7,1.0)	0.7 (0.5,1.0)	0.7 (0.5,1.0)	0.6 (0.4,0.9)	0.8 (0.5, 1.2)	0.8 (0.4, 1.3)	0.8 (0.5,1.2)	0.8 (0.4,1.3)
146. Xylene	0.8 (0.7,1.1)	0.9 (0.5,1.7)	0.8 (0.6,1.1)	0.7 (0.3,1.4)	1.0 (0.6, 1.6)	1.0 (0.5, 1.9)	1.0 (0.6,1.5)	0.9 (0.5,1.8)
147. Styrene	0.4 (0.2,0.7)	0.3 (0.1,0.7)	0.1 (0.1,0.4)		0.5 (0.2, 1.1)		0.5 (0.2,1.1)	
Polycyclic aromatic hydrocarbons [Any, Substantial] [b]								
76. Soot	1.0 (0.7,1.4)	1.6 (0.9,2.7)	1.2 (0.8,1.8)	1.1 (0.6,1.9)	0.9 (0.6, 1.4)	0.9 (0.5, 1.7)	0.9 (0.6,1.4)	1.0 (0.5,1.8)
214. PAH (Any)	1.0 (0.9,1.3)	1.2 (0.8,1.8)	0.9 (0.8,1.1)	1.1 (0.7,1.5)	1.0 (0.7, 1.3)	1.3 (0.8, 2.0)	1.0 (0.7,1.3)	1.2 (0.8,1.9)
219. Benzo(a)pyrene	1.1 (0.9,1.4)	1.0 (0.6,1.6)	1.0 (0.8,1.3)	1.2 (0.7,2.2)	1.1 (0.8, 1.5)	0.9 (0.5, 1.7)	1.1 (0.8,1.5)	1.0 (0.5,1.7)
Engine emissions [Any, Substantial] [b]								
115. Gas Eng.Emissions	0.8 (0.7,1.0)	0.8 (0.7,1.0)	0.9 (0.8,1.1)	1.0 (0.8,1.2)	0.9 (0.6, 1.2)	1.1 (0.8, 1.5)	0.9 (0.7,1.2)	1.1 (0.8,1.5)
117. Diesel Eng.Emissions	1.1 (0.9,1.4)	1.1 (0.9,1.5)	1.1 (0.9,1.4)	1.4 (1.0,1.9)	1.1 (0.9, 1.4)	1.2 (0.8, 1.7)	1.1 (0.9,1.4)	1.2 (0.8,1.7)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
121. Jet Fuel Eng.Emiss.	0.4 (0.1,1.3)		0.5 (0.2,1.6)		0.8 (0.4, 1.8)		0.9 (0.4,1.9)	
122. Propane Eng.Emiss.	1.6 (1.0,2.5)	1.1 (0.6,2.0)	1.5 (0.9,2.5)	1.1 (0.6,2.0)	1.4 (0.8, 2.2)	1.0 (0.5, 1.7)	1.4 (0.8,2.2)	0.9 (0.5,1.7)
Inorganic acid mists [Any, Substantial] [b]								
85. Hydrogen Fluoride	1.7 (1.1,2.6)	1.3 (0.4,4.3)	1.0 (0.6,1.8)	2.2 (0.5,9.0)	1.3 (0.7, 2.1)	1.2 (0.5, 2.7)	1.2 (0.7,2.0)	1.0 (0.4,2.3)
89. Hydrogen Chloride	0.9 (0.7,1.2)	1.0 (0.6,1.6)	1.1 (0.7,1.6)	0.9 (0.5,1.5)	0.9 (0.6, 1.4)	1.1 (0.6, 2.0)	0.9 (0.6,1.4)	1.1 (0.6,1.9)
126. Inorg.Acid Solutions	1.1 (0.9,1.4)	0.9 (0.6,1.4)	1.0 (0.7,1.4)	0.7 (0.4,1.3)	1.1 (0.8, 1.6)	0.7 (0.4, 1.2)	1.1 (0.8,1.6)	0.7 (0.4,1.2)
129. Plating Solutions	0.9 (0.4,1.8)	1.1 (0.4,3.5)	1.0 (0.4,2.3)		0.9 (0.4, 1.8)		0.8 (0.4,1.7)	
130. Nitric Acid	0.8 (0.4,1.5)	0.5 (0.2,1.5)	0.7 (0.3,1.7)	0.5 (0.1,1.8)	0.8 (0.4, 1.5)	0.9 (0.4, 2.0)	0.8 (0.4,1.5)	0.8 (0.3,1.7)
131. Phosphoric Acid	1.5 (0.8,2.9)	0.7 (0.2,3.3)	0.9 (0.4,2.3)		1.1 (0.5, 2.1)		1.0 (0.5,2.0)	
132. Sulphuric Acid	0.9 (0.7,1.2)	0.7 (0.4,1.3)	0.7 (0.5,1.1)	0.7 (0.4,1.3)	0.8 (0.5, 1.2)	0.8 (0.4, 1.6)	0.8 (0.5,1.2)	0.8 (0.4,1.5)
Resins and resin-containing compounds [Any, Substantial] [b]								
53. Natural Rubber	1.2 (0.8,1.7)	0.3 (0.1,0.9)	3.8 (1.9,7.6)		2.3 (1.3, 4.1)		2.1 (1.3,3.5)	
55. Plastic Dust	0.8 (0.6,1.2)	0.7 (0.4,1.4)	0.9 (0.6,1.4)	0.8 (0.4,1.7)	0.9 (0.6, 1.4)	0.9 (0.5, 1.6)	0.9 (0.6,1.4)	0.9 (0.5,1.6)
61. Cellulose Nitrate	0.6 (0.4,1.1)	0.6 (0.3,1.2)	0.9 (0.4,1.8)	0.6 (0.3,1.3)	0.8 (0.4, 1.4)	0.9 (0.4, 1.8)	0.8 (0.5,1.4)	0.9 (0.5,1.7)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
62. Polyvinyl Chloride	0.6 (0.3,1.2)	0.8 (0.2,3.1)	0.4 (0.2,0.9)		0.8 (0.4, 1.4)		0.8 (0.5,1.5)	
63. Polyvinyl Acetate	0.6 (0.4,0.9)		0.5 (0.3,1.0)		0.6 (0.4, 1.1)		0.7 (0.4,1.2)	
64. Poly-Acrylates	1.0 (0.6,1.6)	1.1 (0.5,2.4)	2.9 (1.5,5.7)	1.3 (0.5,3.3)	1.7 (0.9, 3.1)	1.4 (0.7, 2.9)	1.6 (0.9,2.8)	1.3 (0.6,2.5)
65. Alkyds	0.8 (0.6,1.2)	0.8 (0.3,1.7)	1.3 (0.7,2.3)	0.7 (0.3,1.6)	0.9 (0.5, 1.6)	0.8 (0.4, 1.7)	0.9 (0.5,1.6)	0.8 (0.4,1.7)
66. Epoxies	1.6 (0.8,3.1)	2.2 (0.8,6.1)	1.4 (0.6,3.2)	3.3 (0.9,11.8)	1.4 (0.7, 2.6)	1.4 (0.6, 3.2)	1.3 (0.7,2.4)	1.2 (0.6,2.7)
67. Phenol-Formald.	1.3 (0.9,1.8)	0.6 (0.2,1.5)	1.2 (0.7,2.0)	0.6 (0.2,1.8)	1.2 (0.7, 2.0)	0.9 (0.4, 1.9)	1.2 (0.8,2.0)	0.8 (0.4,1.8)
68. Urea-Formald.	1.4 (1.0,1.9)	0.7 (0.2,1.9)	2.0 (1.2,3.5)	0.8 (0.3,2.4)	1.7 (1.0, 2.7)	1.0 (0.5, 2.2)	1.6 (1.0,2.5)	1.0 (0.5,2.0)
69. Polyurethanes	1.2 (0.7,2.1)	1.4 (0.5,3.5)	1.8 (0.9,3.6)	2.0 (0.7,5.6)	1.2 (0.6, 2.2)	1.0 (0.5, 2.3)	1.2 (0.7,2.1)	1.0 (0.5,2.1)
70. Styrene-Buta.Rubber	0.8 (0.6,1.2)	0.2 (0.1,0.8)	0.3 (0.2,0.7)		0.6 (0.3, 1.0)		0.6 (0.4,1.1)	
71. Polychloroprene	1.0 (0.7,1.5)	1.2 (0.3,4.3)	1.0 (0.6,1.7)		1.0 (0.6, 1.8)		1.1 (0.6,1.7)	
149. Animal & Vege.Glues	0.9 (0.6,1.3)	0.9 (0.5,1.6)	0.8 (0.5,1.3)	0.9 (0.5,1.7)	0.9 (0.6, 1.3)	0.9 (0.5, 1.5)	0.9 (0.6,1.3)	0.8 (0.5,1.4)
151. Linseed Oil	1.1 (0.8,1.6)	0.9 (0.3,2.4)	0.9 (0.6,1.4)	0.7 (0.2,1.9)	0.9 (0.6, 1.5)	1.0 (0.4, 2.2)	0.9 (0.6,1.5)	0.9 (0.5,2.0)
152. Synthetic Adhesives	0.9 (0.8,1.2)	0.8 (0.6,1.1)	0.8 (0.6,1.1)	0.7 (0.5,0.9)	0.8 (0.6, 1.1)	0.7 (0.5, 1.0)	0.8 (0.6,1.1)	0.7 (0.5,1.0)
173. Wood Varnishes, Stains	1.1 (0.8,1.5)	1.3 (0.8,1.9)	1.5 (1.0,2.4)	1.6 (1.0,2.6)	1.1 (0.7, 1.7)	1.3 (0.8, 2.2)	1.1 (0.7,1.7)	1.2 (0.7,2.1)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
174. Inks	1.5 (1.1,2.3)	1.1 (0.7,1.9)	1.8 (1.1,2.7)	0.9 (0.5,1.7)	1.3 (0.8, 2.1)	1.0 (0.5, 1.8)	1.3 (0.8,2.1)	1.0 (0.5,1.8)
175. Metal Coatings	1.1 (0.8,1.4)	1.0 (0.6,1.5)	1.4 (1.0,2.0)	1.0 (0.6,1.6)	1.3 (0.9, 1.8)	1.1 (0.7, 1.9)	1.2 (0.9,1.8)	1.1 (0.7,1.8)
Carbonaceous compounds	[Any, Substantial] [b]							
73. Coal Dust	1.2 (0.9,1.7)	1.4 (0.9,2.2)	1.0 (0.7,1.4)	1.2 (0.7,1.9)	1.0 (0.7, 1.5)	1.1 (0.7, 1.9)	1.0 (0.7,1.5)	1.1 (0.6,1.9)
74. Carbon Black	1.2 (0.8,1.7)	0.7 (0.3,1.6)	1.0 (0.6,1.7)	0.7 (0.2,1.9)	1.0 (0.6, 1.6)	0.9 (0.4, 1.9)	0.9 (0.6,1.5)	0.9 (0.3,2.2)
78. Graphite Dust	0.6 (0.3,1.2)	0.4 (0.1,2.4)	0.3 (0.1,0.8)		0.6 (0.3, 1.2)		0.6 (0.3,1.2)	
Aliphatic alkanes (C5-C17)	[Any, Substantial] [b]							
155. Leaded Gasoline	1.1 (0.8,1.3)	0.9 (0.6,1.3)	1.1 (0.9,1.5)	0.8 (0.6,1.2)	1.0 (0.7, 1.4)	0.9 (0.5, 1.4)	1.0 (0.7,1.5)	0.9 (0.6,1.6)
156. Kerosene	1.6 (1.2,2.1)	2.6 (1.5,4.4)	1.4 (1.0,2.0)	2.6 (1.5,4.6)	1.2 (0.9, 1.8)	1.8 (1.0, 3.0)	1.2 (0.9,1.7)	1.8 (1.1,3.1)
157. Diesel Oil	1.2 (0.8,1.8)	1.3 (0.7,2.4)	1.6 (1.1,2.4)	1.7 (0.9,3.2)	1.2 (0.8, 1.8)	1.2 (0.7, 2.2)	1.2 (0.8,1.8)	1.4 (0.8,2.4)
158. Heating Oil	1.3 (0.9,1.9)	1.9 (1.1,3.3)	1.2 (0.8,1.7)	2.0 (1.1,3.5)	1.2 (0.8, 1.9)	1.5 (0.9, 2.6)	1.2 (0.8,1.8)	1.6 (0.9,2.7)
159. Mineral Spirits	1.1 (0.9,1.4)	1.2 (0.9,1.7)	1.3 (1.0,1.7)	1.3 (1.0,1.9)	1.1 (0.8, 1.5)	1.1 (0.8, 1.6)	1.1 (0.8,1.5)	1.2 (0.8,1.7)
167. Jet Fuel	0.6 (0.3,1.4)	0.5 (0.2,1.4)	0.4 (0.2,1.0)		0.7 (0.3, 1.5)		0.8 (0.4,1.5)	
168. Aviation Gasoline	0.5 (0.2,1.2)	0.5 (0.2,1.3)	0.4 (0.2,1.0)		0.6 (0.3, 1.4)		0.7 (0.4,1.4)	

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
Aliphatic alcohols [Any, Substantial] [b]								
133. Methanol	0.8 (0.6,1.2)	1.3 (0.7,2.3)	1.0 (0.7,1.6)	1.5 (0.8,2.9)	0.8 (0.5, 1.4)	1.1 (0.6, 2.1)	0.9 (0.5,1.3)	1.2 (0.7,2.4)
134. Ethanol	1.5 (0.8,2.6)	1.2 (0.3,5.1)	1.9 (1.0,3.6)	1.8 (0.4,8.1)	1.4 (0.7, 2.5)	1.2 (0.5, 2.7)	1.3 (0.7,2.3)	1.4 (0.5,3.4)
135. Ethylene Glycol	0.7 (0.5,1.0)		1.1 (0.7,1.9)		1.0 (0.6, 1.8)		1.0 (0.6,1.7)	
136. Isopropanol	1.1 (0.7,1.5)	1.2 (0.7,2.2)	1.2 (0.7,1.8)	1.2 (0.6,2.3)	0.9 (0.6, 1.5)	1.2 (0.7, 2.2)	0.9 (0.6,1.5)	1.3 (0.7,2.5)
Aliphatic chlorinated hydrocarbons [Any, Substantial] [b]								
138. Carbon Tetrachloride	1.0 (0.7,1.4)	1.5 (0.9,2.5)	1.4 (0.9,2.2)	1.6 (0.9,2.8)	1.1 (0.7, 1.8)	1.4 (0.8, 2.4)	1.1 (0.7,1.7)	1.6 (0.9,2.8)
139. Methylene Chloride	0.8 (0.5,1.3)	1.8 (0.8,4.2)	1.1 (0.5,2.1)	2.0 (0.8,4.9)	0.9 (0.5, 1.6)	1.1 (0.5, 2.3)	0.9 (0.5,1.7)	1.4 (0.6,3.2)
140. 1,1,1.-Trichlorethane	1.7 (0.9,3.1)	1.8 (0.8,3.9)	2.7 (1.3,5.7)	2.2 (0.9,5.4)	1.4 (0.8, 2.6)	1.4 (0.7, 2.8)	1.4 (0.8,2.4)	1.6 (0.8,3.4)
141. Trichloroethylene	1.2 (0.8,1.9)	0.8 (0.4,1.5)	1.2 (0.7,2.1)		1.1 (0.6, 2.1)		1.1 (0.6,2.0)	
142. Perchloroethylene	1.0 (0.5,1.9)	1.0 (0.4,2.2)	1.1 (0.5,2.3)		0.9 (0.5, 1.7)		0.9 (0.5,1.8)	
209. Chlorinated Alkenes	1.1 (0.8,1.6)	0.9 (0.5,1.6)	1.2 (0.8,1.8)		1.1 (0.6, 1.9)		1.1 (0.6,1.8)	
Inorganic gases [Any, Substantial] [b]								
79. Hydrogen	0.8 (0.5,1.3)	1.1 (0.3,3.8)	1.3 (0.6,2.8)		0.9 (0.5, 1.7)		0.9 (0.5,1.6)	

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
80. Carbon Monoxide	1.1 (0.9,1.3)	1.0 (0.8,1.4)	0.9 (0.8,1.1)	0.8 (0.5,1.2)	1.1 (0.8, 1.5)	0.9 (0.5, 1.5)	1.1 (0.8,1.5)	0.9 (0.5,1.4)
81. Hydrogen Cyanide	0.9 (0.5,1.6)	0.6 (0.3,1.5)	2.0 (0.8,4.9)		1.3 (0.6, 2.6)		1.2 (0.6,2.4)	
82. Ammonia	0.9 (0.7,1.1)	0.8 (0.5,1.1)	1.0 (0.7,1.4)	0.8 (0.5,1.3)	1.1 (0.8, 1.5)	0.9 (0.5, 1.4)	1.1 (0.8,1.5)	0.9 (0.5,1.4)
83. Nitrogen Oxides	1.6 (1.3,2.0)	1.6 (1.0,2.5)	2.6 (1.9,3.7)	2.4 (1.4,4.2)	2.1 (1.5, 3.1)	1.7 (1.0, 3.0)	2.1 (1.5,2.9)	1.6 (1.0,2.8)
84. Ozone	1.3 (0.9,1.8)	0.9 (0.4,2.2)	0.8 (0.5,1.3)	0.4 (0.1,1.5)	0.9 (0.6, 1.5)	0.7 (0.3, 1.6)	0.9 (0.6,1.5)	0.8 (0.4,1.6)
85. Hydrogen Fluoride	1.7 (1.1,2.6)	1.3 (0.4,4.3)	1.0 (0.6,1.8)	2.2 (0.5,9.0)	1.3 (0.7, 2.1)	1.2 (0.5, 2.7)	1.2 (0.7,2.0)	1.0 (0.4,2.3)
86. Sulphur Dioxide	0.9 (0.7,1.1)	0.7 (0.4,1.4)	0.5 (0.4,0.7)	0.4 (0.2,0.9)	0.6 (0.4, 0.8)	0.8 (0.4, 1.6)	0.6 (0.4,0.9)	0.9 (0.5,1.6)
87. Hydrogen Sulphide	1.0 (0.7,1.4)	0.6 (0.3,1.5)	1.5 (0.9,2.3)	1.4 (0.5,3.7)	1.3 (0.8, 2.0)	1.1 (0.5, 2.4)	1.3 (0.8,2.0)	1.1 (0.5,2.3)
88. Chlorine	0.5 (0.3,0.8)	0.7 (0.3,1.5)	0.6 (0.3,1.2)	1.1 (0.4,2.9)	0.7 (0.4, 1.3)	1.0 (0.5, 2.1)	0.7 (0.4,1.3)	1.0 (0.5,2.0)
89. Hydrogen Chloride	0.9 (0.7,1.2)	1.0 (0.6,1.6)	1.1 (0.7,1.6)	0.9 (0.5,1.5)	0.9 (0.6, 1.4)	1.1 (0.6, 2.0)	0.9 (0.6,1.4)	1.1 (0.6,1.9)
97. Coal Gas	0.6 (0.3,1.1)	1.4 (0.3,6.8)	0.5 (0.2,1.2)		0.7 (0.3, 1.5)		0.7 (0.4,1.4)	
Organic gases (C1-C4) [Any, Substantial] [b]								
91. Methane	0.9 (0.6,1.3)	2.6 (1.3,5.4)	1.0 (0.7,1.6)	2.5 (1.2,5.5)	1.1 (0.7, 1.7)	1.7 (0.8, 3.2)	1.0 (0.6,1.7)	2.1 (0.9,4.9)
92. Propane	1.1 (0.8,1.6)	0.8 (0.2,3.4)	1.8 (0.7,4.7)	0.9 (0.2,4.3)	1.2 (0.6, 2.1)	0.9 (0.4, 2.2)	1.1 (0.6,2.0)	1.5 (0.5,4.5)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
94. Acetylene	1.3 (0.9,1.9)	0.3 (0.0,2.4)	0.9 (0.6,1.4)		0.9 (0.6, 1.4)		0.9 (0.6,1.5)	
Inorganic salts [Any, Substantial] [b]								
25. Sodium Carbonate	1.0 (0.5,1.9)		0.8 (0.3,1.7)		1.1 (0.6, 2.1)		1.1 (0.5,2.1)	
30. Calcium Sulphate	1.0 (0.8,1.3)	1.2 (0.9,1.7)	1.1 (0.8,1.5)	1.3 (0.9,1.9)	1.0 (0.7, 1.5)	1.2 (0.8, 1.8)	1.0 (0.7,1.5)	1.2 (0.8,1.9)
31. Calcium Carbonate	0.8 (0.6,1.2)	1.2 (0.6,2.2)	0.8 (0.5,1.2)	1.3 (0.7,2.6)	0.9 (0.6, 1.3)	1.3 (0.7, 2.5)	0.9 (0.6,1.3)	1.4 (0.7,2.6)
Magnesium compounds [Any] [b]								
10. Industrial Talc	0.7 (0.5,1.1)	0.6 (0.3,1.3)	0.9 (0.5,1.5)	0.6 (0.2,1.3)	0.8 (0.5, 1.3)	0.8 (0.4, 1.7)	0.8 (0.5,1.4)	0.9 (0.5,1.8)
23. Cosmetic Talc	1.5 (0.8,2.7)	0.4 (0.1,1.6)	1.9 (1.0,3.7)		1.6 (0.9, 3.0)		1.6 (0.8,3.2)	
Aluminium compounds [Any, Substantial] [b]								
12. Clay Dust	1.6 (1.0,2.6)	1.1 (0.5,2.4)	2.1 (1.2,3.6)	1.2 (0.5,3.0)	1.6 (1.0, 2.8)	1.2 (0.6, 2.4)	1.8 (1.0,3.2)	1.2 (0.6,2.7)
21. Aluminium Alloy Dust	1.4 (1.0,1.9)	1.1 (0.7,1.8)	1.5 (1.0,2.3)	1.6 (0.9,2.6)	1.3 (0.9, 1.9)	1.2 (0.7, 2.1)	1.3 (0.9,1.9)	1.2 (0.7,2.2)
26. Alumina	1.2 (0.9,1.5)	1.4 (0.9,2.2)	1.6 (1.2,2.3)	1.7 (0.9,3.0)	1.4 (1.0, 1.9)	1.6 (0.9, 2.8)	1.4 (1.0,1.9)	1.7 (1.0,3.1)
102. Aluminium Fumes	1.2 (0.7,1.9)	0.8 (0.4,1.9)	0.6 (0.3,1.2)	0.3 (0.1,0.7)	0.7 (0.4, 1.2)	0.7 (0.3, 1.4)	0.7 (0.4,1.3)	0.7 (0.3,1.6)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
Chromates [Any] [b]								
40. Lead Chromate	0.7 (0.4,1.2)	1.1 (0.2,5.6)	1.3 (0.7,2.6)		1.1 (0.6, 2.0)		1.4 (0.7,2.8)	
104. Chromium Fumes	2.0 (1.2,3.2)	2.1 (0.7,6.6)	1.3 (0.8,2.2)	1.7 (0.6,5.1)	1.4 (0.7, 2.9)	1.4 (0.7, 3.1)	1.9 (0.6,6.1)	1.6 (0.7,4.0)
Manganese compounds [Any, Substantial] [b]								
16. Stainless Steel Dust	1.4 (1.0,1.9)	1.5 (0.9,2.6)	1.2 (0.8,1.9)	1.5 (0.8,2.7)	1.1 (0.7, 1.6)	1.3 (0.7, 2.4)	1.1 (0.7,1.7)	1.4 (0.7,2.7)
17. Mild Steel Dust	1.2 (1.0,1.5)	1.1 (0.8,1.5)	1.2 (0.9,1.6)	1.2 (0.8,1.7)	1.2 (0.9, 1.6)	1.1 (0.7, 1.6)	1.2 (0.9,1.6)	1.2 (0.8,1.8)
105. Manganese Fumes	1.4 (1.0,1.9)	3.3 (1.5,7.0)	1.0 (0.6,1.6)	1.8 (0.7,4.6)	1.1 (0.6, 1.8)	1.4 (0.7, 2.9)	1.2 (0.7,2.2)	2.2 (0.9,5.5)
Iron compounds [Any, Substantial] [b]								
16. Stainless Steel Dust	1.4 (1.0,1.9)	1.5 (0.9,2.6)	1.2 (0.8,1.9)	1.5 (0.8,2.7)	1.1 (0.7, 1.6)	1.3 (0.7, 2.4)	1.1 (0.7,1.7)	1.4 (0.7,2.7)
17. Mild Steel Dust	1.2 (1.0,1.5)	1.1 (0.8,1.5)	1.2 (0.9,1.6)	1.2 (0.8,1.7)	1.2 (0.9, 1.6)	1.1 (0.7, 1.6)	1.2 (0.9,1.6)	1.2 (0.8,1.8)
33. Iron Dust	0.9 (0.6,1.4)	0.6 (0.3,1.2)	0.9 (0.5,1.4)	0.5 (0.2,1.0)	0.9 (0.6, 1.4)	0.7 (0.4, 1.4)	0.8 (0.5,1.3)	0.6 (0.3,1.1)
34. Iron Oxides	0.9 (0.7,1.1)	0.7 (0.5,1.1)	0.8 (0.6,1.1)	0.7 (0.4,1.1)	0.8 (0.5, 1.1)	0.8 (0.5, 1.3)	0.7 (0.5,1.0)	0.8 (0.5,1.2)
106. Iron Fumes	1.2 (0.9,1.6)	1.4 (1.0,2.1)	0.9 (0.6,1.3)	0.8 (0.5,1.5)	0.8 (0.5, 1.3)	1.0 (0.5, 1.9)	0.7 (0.4,1.2)	0.6 (0.3,1.3)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
Nickel compounds [Any] [b]								
16. Stainless Steel Dust	1.4 (1.0,1.9)	1.5 (0.9,2.6)	1.2 (0.8,1.9)	1.5 (0.8,2.7)	1.1 (0.7, 1.6)	1.3 (0.7, 2.4)	1.1 (0.7,1.7)	1.4 (0.7,2.7)
107. Nickel Fumes	1.7 (1.1,2.8)		1.3 (0.8,2.1)		1.1 (0.5, 2.3)		0.8 (0.3,2.5)	
Copper compounds [Any, Substantial] [b]								
14. Bronze Dust	0.8 (0.4,1.6)	0.7 (0.3,1.9)	0.6 (0.3,1.4)	0.5 (0.2,1.4)	0.8 (0.4, 1.4)	0.8 (0.4, 1.7)	0.7 (0.4,1.5)	0.6 (0.2,1.8)
15. Brass Dust	1.4 (0.9,2.3)	1.5 (0.7,3.1)	1.2 (0.7,2.1)	1.3 (0.6,3.0)	1.2 (0.7, 2.1)	1.3 (0.7, 2.6)	1.3 (0.8,2.3)	2.0 (0.9,4.5)
35. Copper Dust	1.2 (0.8,1.7)	2.6 (1.3,5.3)	1.2 (0.7,1.9)	2.1 (1.0,4.5)	1.2 (0.7, 1.9)	1.7 (0.9, 3.2)	1.2 (0.8,2.0)	1.9 (0.9,3.9)
108. Copper Fumes	1.8 (1.2,2.7)	2.1 (1.1,3.9)	1.5 (0.9,2.4)	2.2 (1.0,4.5)	1.5 (0.9, 2.6)	1.3 (0.7, 2.6)	1.5 (0.9,2.7)	1.6 (0.7,3.6)
Zinc compounds [Any, Substantial] [b]								
15. Brass Dust	1.4 (0.9,2.3)	1.5 (0.7,3.1)	1.2 (0.7,2.1)	1.3 (0.6,3.0)	1.2 (0.7, 2.1)	1.3 (0.7, 2.6)	1.3 (0.8,2.3)	2.0 (0.9,4.5)
36. Zinc Dust	1.4 (0.9,2.3)	3.5 (1.1,11.1)	1.6 (0.9,2.9)	2.7 (0.7,10.7)	1.2 (0.7, 2.1)	1.3 (0.6, 3.1)	1.2 (0.7,2.1)	1.7 (0.6,4.9)
37. Zinc Oxide	1.0 (0.7,1.6)	0.8 (0.2,2.8)	1.5 (0.9,2.7)		1.3 (0.7, 2.2)		1.3 (0.7,2.3)	
109. Zinc Fumes	1.3 (0.9,1.9)	2.6 (1.3,5.1)	0.9 (0.5,1.6)	1.5 (0.6,3.7)	0.9 (0.5, 1.5)	1.2 (0.6, 2.4)	0.9 (0.5,1.5)	1.4 (0.6,3.3)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
Tin compounds [Any, Substantial] [b]								
14. Bronze Dust	0.8 (0.4,1.6)	0.7 (0.3,1.9)	0.6 (0.3,1.4)	0.5 (0.2,1.4)	0.8 (0.4, 1.4)	0.8 (0.4, 1.7)	0.7 (0.4,1.5)	0.6 (0.2,1.8)
111. Tin Fumes	1.4 (1.0,2.0)	1.0 (0.5,2.1)	1.1 (0.6,2.1)	0.5 (0.2,1.3)	1.0 (0.6, 1.8)	1.0 (0.5, 2.0)	0.8 (0.4,1.6)	0.6 (0.2,1.7)
Lead compounds [Any, Substantial] [b]								
38. Lead Oxides	1.5 (0.9,2.5)	1.9 (0.8,4.9)	1.3 (0.7,2.5)	1.7 (0.7,4.7)	1.3 (0.8, 2.3)	1.4 (0.7, 3.0)	1.3 (0.7,2.4)	1.6 (0.6,3.8)
39. Basic Lead Carb.	1.2 (0.7,1.9)	1.4 (0.3,6.1)	1.7 (1.0,3.0)		1.2 (0.6, 2.2)		1.2 (0.6,2.2)	
40. Lead Chromate	0.7 (0.4,1.2)	1.1 (0.2,5.6)	1.3 (0.7,2.6)		1.1 (0.6, 2.0)		1.4 (0.7,2.8)	
112. Lead Fumes	1.3 (0.9,1.8)	0.8 (0.3,1.6)	1.0 (0.5,2.0)	0.5 (0.2,1.1)	1.1 (0.6, 1.9)	0.8 (0.4, 1.8)	1.2 (0.6,2.1)	0.9 (0.4,2.1)
155. Leaded Gasoline	1.1 (0.8,1.3)	0.9 (0.6,1.3)	1.1 (0.9,1.5)	0.8 (0.6,1.2)	1.0 (0.7, 1.4)	0.9 (0.5, 1.4)	1.0 (0.7,1.5)	0.9 (0.6,1.6)
Not in a category								
1. Abrasives Dust	1.1 (0.9,1.3)	0.9 (0.7,1.2)	0.7 (0.5,1.0)	0.8 (0.6,1.1)	0.8 (0.6, 1.0)	0.8 (0.5, 1.1)	0.8 (0.6,1.0)	0.8 (0.5,1.1)
18. Inorg.Pigments	1.2 (0.9,1.5)	1.4 (0.8,2.4)	1.5 (1.0,2.2)	3.0 (1.3,7.0)	1.4 (0.9, 2.1)	1.3 (0.7, 2.5)	1.3 (0.9,2.0)	1.2 (0.7,2.3)
20. Extenders	0.8 (0.6,1.1)	0.7 (0.3,1.3)	0.8 (0.5,1.3)	0.3 (0.1,0.7)	0.8 (0.5, 1.3)	0.7 (0.4, 1.5)	0.8 (0.5,1.3)	0.8 (0.4,1.5)
22. Ashes	1.3 (0.8,2.1)	1.5 (0.8,3.0)	2.3 (1.3,4.1)	2.2 (1.0,4.8)	1.5 (0.8, 2.7)	1.4 (0.7, 2.9)	1.5 (0.8,2.5)	1.3 (0.7,2.5)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
24. Borates	1.7 (0.9,3.5)	4.1 (1.1,15.9)	2.0 (0.8,5.4)		1.3 (0.6, 2.6)		1.2 (0.7,2.3)	
28. Sulfur	0.7 (0.4,1.4)	1.4 (0.4,5.4)	0.7 (0.3,1.5)	2.5 (0.6,10.3)	0.7 (0.4, 1.4)	1.3 (0.5, 3.0)	0.8 (0.4,1.4)	1.2 (0.6,2.4)
41. Organic Dyes & Pig.	0.9 (0.7,1.2)	1.3 (0.7,2.5)	0.8 (0.5,1.2)	1.6 (0.9,3.2)	0.9 (0.6, 1.3)	1.3 (0.7, 2.4)	0.9 (0.6,1.3)	1.2 (0.7,2.1)
56. Rayon Fibres	0.9 (0.6,1.5)	1.3 (0.6,2.6)	1.0 (0.5,1.8)	1.5 (0.6,3.5)	0.9 (0.5, 1.6)	1.3 (0.7, 2.8)	0.9 (0.5,1.5)	1.3 (0.7,2.4)
57. Acrylic Fibres	0.9 (0.5,1.5)	1.1 (0.5,2.2)	0.7 (0.3,1.4)	1.2 (0.6,2.5)	0.9 (0.5, 1.6)	1.1 (0.6, 2.3)	0.9 (0.5,1.5)	1.1 (0.6,2.1)
58. Polyester Fibres	1.1 (0.7,1.6)	0.9 (0.5,1.5)	1.2 (0.7,2.0)	0.9 (0.5,1.6)	1.1 (0.7, 2.0)	1.2 (0.6, 2.3)	1.1 (0.7,1.9)	1.1 (0.6,2.0)
59. Nylon Fibres	1.1 (0.7,1.7)	0.9 (0.5,1.8)	1.5 (0.9,2.6)	1.0 (0.5,2.1)	1.3 (0.8, 2.3)	1.1 (0.5, 2.1)	1.3 (0.7,2.1)	1.1 (0.6,2.0)
60. Acetate Fibres	0.8 (0.4,1.5)	0.8 (0.2,2.5)	0.6 (0.3,1.3)	1.0 (0.3,3.2)	0.9 (0.4, 1.6)	0.8 (0.4, 1.9)	0.9 (0.5,1.6)	0.8 (0.4,1.7)
93. Formaldehyde	0.9 (0.7,1.1)	0.8 (0.5,1.2)	0.9 (0.7,1.2)	0.8 (0.5,1.4)	1.0 (0.8, 1.4)	0.9 (0.5, 1.6)	1.0 (0.8,1.3)	0.9 (0.6,1.5)
95. Phosgene	0.7 (0.3,1.3)	5.3 (0.4,73.6)	1.2 (0.5,3.1)		0.9 (0.5, 1.9)		0.9 (0.5,1.8)	
96. Spray Gases	1.0 (0.6,1.8)	0.6 (0.2,1.7)	0.8 (0.4,1.5)	0.5 (0.2,1.6)	1.0 (0.5, 1.8)	0.8 (0.4, 1.7)	1.0 (0.5,1.7)	0.8 (0.4,1.6)
114. Cooking Fumes	0.8 (0.6,1.1)	0.9 (0.6,1.3)	0.7 (0.5,1.0)	0.8 (0.5,1.3)	0.8 (0.6, 1.2)	0.9 (0.6, 1.5)	0.8 (0.6,1.2)	0.9 (0.6,1.4)
116. Coal Comb.Products	1.2 (0.8,1.7)	0.9 (0.5,1.8)	0.9 (0.6,1.4)	0.9 (0.5,1.6)	1.0 (0.7, 1.6)	0.7 (0.4, 1.5)	1.0 (0.7,1.6)	0.8 (0.4,1.5)
118. Liquid Fuel Comb.Prod.	1.1 (0.8,1.5)	0.9 (0.6,1.4)	0.9 (0.6,1.3)	0.6 (0.4,1.1)	0.9 (0.6, 1.4)	0.8 (0.5, 1.4)	1.0 (0.7,1.4)	0.8 (0.5,1.3)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using P<0.25. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
120. Natural Gas Comb.Prod.	0.7 (0.5,1.1)	0.9 (0.4,2.1)	0.5 (0.3,0.8)	0.5 (0.2,1.2)	0.7 (0.4, 1.1)	0.8 (0.4, 1.6)	0.7 (0.4,1.1)	0.8 (0.4,1.5)
123. Plastics Pyrol.Prod.	0.5 (0.3,0.8)	0.4 (0.2,0.9)	0.8 (0.4,1.6)	0.5 (0.2,1.1)	0.7 (0.4, 1.3)	0.7 (0.3, 1.5)	0.7 (0.4,1.3)	0.8 (0.4,1.5)
124. Rubber Pyrol.Prod.	0.9 (0.5,1.5)	0.7 (0.3,1.5)	1.2 (0.6,2.5)	1.6 (0.6,4.2)	1.0 (0.5, 1.9)	1.3 (0.6, 2.8)	1.0 (0.6,1.8)	1.2 (0.6,2.4)
125. Propane Comb.Prod.	1.0 (0.6,1.4)	1.6 (0.8,3.1)	0.6 (0.4,1.0)	1.5 (0.7,3.3)	0.7 (0.3, 1.2)	0.8 (0.4, 1.8)	0.7 (0.4,1.3)	0.8 (0.4,1.7)
127. Alkali, Caustic Solutions	1.3 (0.9,1.7)	1.5 (0.9,2.5)	1.7 (1.2,2.4)	2.1 (1.3,3.6)	1.6 (1.1, 2.2)	1.7 (1.0, 2.8)	1.5 (1.1,2.2)	1.6 (1.0,2.6)
137. Acetic Acid	1.0 (0.6,1.6)	1.1 (0.5,2.3)	1.0 (0.6,1.6)	1.7 (0.8,3.8)	1.1 (0.7, 1.9)	1.3 (0.7, 2.6)	1.1 (0.7,1.8)	1.2 (0.7,2.2)
143. Acetone	1.0 (0.6,1.6)	1.2 (0.6,2.4)	0.8 (0.4,1.5)	1.1 (0.5,2.3)	0.9 (0.5, 1.5)	1.1 (0.6, 2.1)	0.9 (0.5,1.5)	1.1 (0.6,2.0)
148. Phenol	0.7 (0.4,1.4)		0.8 (0.3,2.1)		0.8 (0.4, 1.6)		0.8 (0.5,1.6)	
150. Turpentine	1.0 (0.7,1.4)	1.2 (0.8,1.8)	1.2 (0.8,1.7)	1.2 (0.8,1.9)	0.9 (0.6, 1.5)	0.8 (0.5, 1.4)	0.9 (0.6,1.4)	0.8 (0.5,1.4)
153. Solvents	1.1 (0.9,1.3)	1.1 (0.9,1.4)	1.1 (0.9,1.4)	1.2 (1.0,1.6)	1.1 (0.9, 1.4)	1.2 (0.9, 1.6)	1.1 (0.9,1.4)	1.1 (0.9,1.5)
154. Waxes, Polishes	0.9 (0.7,1.2)	0.9 (0.5,1.5)	1.1 (0.7,1.6)	1.0 (0.5,1.7)	1.0 (0.7, 1.6)	0.9 (0.5, 1.6)	1.1 (0.7,1.5)	0.9 (0.6,1.5)
160. Lubric.Oils & Greases	1.1 (0.9,1.4)	1.1 (0.8,1.5)	1.1 (0.9,1.4)	1.1 (0.8,1.5)	1.0 (0.8, 1.3)	1.1 (0.8, 1.6)	1.0 (0.8,1.3)	1.1 (0.8,1.6)
161. Cutting Fluids	1.3 (1.0,1.7)	1.2 (0.8,1.7)	1.4 (1.0,2.0)	1.0 (0.6,1.5)	1.2 (0.8, 1.7)	1.0 (0.6, 1.6)	1.2 (0.8,1.7)	0.9 (0.6,1.5)
162. Asphalt	0.8 (0.5,1.2)	0.9 (0.5,1.7)	0.7 (0.4,1.1)	0.7 (0.3,1.3)	0.8 (0.5, 1.3)	1.0 (0.5, 1.8)	0.8 (0.5,1.3)	0.9 (0.5,1.7)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
164. Creosote	0.7 (0.3,1.6)	0.8 (0.2,3.1)	0.5 (0.2,1.4)	0.6 (0.1,2.6)	0.8 (0.4, 1.7)	0.8 (0.4, 1.9)	0.8 (0.4,1.6)	0.8 (0.4,1.7)
165. Hydraulic Fluid	0.8 (0.6,1.3)	0.9 (0.3,2.9)	1.1 (0.7,1.8)	0.8 (0.2,2.6)	1.0 (0.6, 1.6)	1.0 (0.5, 2.2)	1.0 (0.6,1.6)	1.0 (0.5,2.0)
179. Hypochlorites	0.7 (0.5,1.0)	0.9 (0.6,1.3)	0.7 (0.5,1.1)	0.9 (0.6,1.4)	0.9 (0.6, 1.4)	1.0 (0.6, 1.7)	0.9 (0.6,1.3)	1.0 (0.6,1.6)
180. Nitrates	0.9 (0.4,1.9)	0.6 (0.1,2.3)	0.3 (0.1,0.8)	0.2 (0.1,1.1)	0.6 (0.3, 1.3)	0.7 (0.3, 1.6)	0.7 (0.4,1.3)	0.7 (0.4,1.5)
185. Vanadium Compounds	1.3 (0.6,2.9)		0.8 (0.4,1.8)		1.0 (0.5, 2.1)		1.0 (0.5,2.0)	
189. Cobalt Compounds	1.3 (0.7,2.1)	0.8 (0.3,2.1)	1.1 (0.6,2.1)		1.2 (0.6, 2.1)		1.2 (0.7,2.0)	
197. Antimony Compounds	1.3 (0.8,2.3)	1.1 (0.4,3.2)	0.8 (0.4,1.6)	1.2 (0.4,3.9)	0.9 (0.5, 1.6)	1.0 (0.5, 2.3)	0.9 (0.5,1.6)	1.0 (0.5,2.0)
198. Tungsten Compounds	1.4 (0.7,2.6)	0.9 (0.4,2.3)	1.3 (0.6,2.8)		1.1 (0.5, 2.2)		1.1 (0.6,2.0)	
199. Gold Compounds	1.8 (0.9,3.5)	2.2 (0.7,7.2)	1.5 (0.6,3.8)		1.3 (0.6, 2.5)		1.2 (0.7,2.3)	
200. Mercury Compounds	1.0 (0.6,1.8)		1.3 (0.6,3.0)		1.3 (0.7, 2.4)		1.2 (0.7,2.2)	
212. Fluorocarbons	0.6 (0.3,1.0)	0.2 (0.0,1.2)	0.5 (0.3,1.0)		0.6 (0.3, 1.1)		0.6 (0.4,1.2)	
213. Glycol Ethers	1.0 (0.6,1.5)	3.3 (1.2,8.9)	1.2 (0.7,2.1)	3.6 (1.1,11.3)	1.0 (0.6, 1.7)	1.6 (0.7, 3.5)	1.0 (0.6,1.6)	1.4 (0.7,2.8)
223. Phthalates	0.4 (0.2,0.7)	1.0 (0.3,3.3)	0.4 (0.2,0.7)		0.6 (0.3, 1.2)		0.6 (0.4,1.2)	
224. Isocyanates	0.9 (0.5,1.6)	0.8 (0.2,3.5)	2.7 (1.1,6.4)		1.3 (0.6, 2.5)		1.2 (0.6,2.3)	

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Table 7-11: Rate ratio estimates, from strategies 3, 4, 7, and 8, for 184 chemicals, at two levels of exposure, grouped by the categories of exchangeability used in the semi-Bayes modeling

	One-chemical-at-a-time regression models				Single large model with simultaneous adjustment			
	Strategy 3 [a]		Strategy 4 [a]		Strategy 7 [a]		Strategy 8 [a]	
	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)	Any RR (90% CL)	Substantial RR (90% CL)
228. Fertilizers	1.3 (0.9,1.8)	1.0 (0.6,1.6)	1.5 (1.0,2.3)	1.1 (0.6,1.8)	1.4 (0.9, 2.2)	1.1 (0.6, 1.8)	1.4 (0.9,2.1)	1.0 (0.6,1.7)
229. Pesticides	1.1 (0.8,1.5)	0.8 (0.5,1.5)	1.1 (0.8,1.7)	0.9 (0.5,1.4)	1.0 (0.6, 1.6)	1.0 (0.6, 1.8)	1.0 (0.7,1.6)	1.0 (0.6,1.7)
230. Biocides	0.8 (0.6,1.0)	1.0 (0.7,1.4)	0.7 (0.5,1.0)	0.9 (0.6,1.5)	0.8 (0.6, 1.2)	1.0 (0.6, 1.7)	0.8 (0.6,1.2)	1.0 (0.6,1.6)
231. Bleaches	0.5 (0.2,1.1)	0.7 (0.2,2.3)	0.5 (0.2,1.3)	0.7 (0.2,2.6)	0.7 (0.4, 1.5)	0.9 (0.4, 2.0)	0.8 (0.4,1.5)	0.9 (0.4,1.8)

[a] All strategies include adjustment for eight non-occupational confounders (see text). Each chemical assessed in a separate regression model: strategy 3, adjustment for seven currently suspected lung carcinogens, strategy 4, automatic forward selection of other chemicals as confounders using $P<0.25$. All chemicals assessed in a single large regression model: strategy 7, semi-Bayes shrinkage of estimates from strategy 6 using a common prior, strategy 8, semi-Bayes shrinkage of estimates from strategy 6 using sets of exchangeability. [b] Not all of the categories of exchangeability were used in the modeling at the substantial level of exposure, so a descriptor follows each category label, indicating in which exposure-level analyses it was used.

Many estimates were consistent across all approaches. The first chemical listed, asbestos (5) for instance, had estimates close to 1.0 at any level of exposure across all models, and estimates close to 1.5 at the substantial level of exposure in three of the four models. Taking into account the unadjusted estimates for asbestos from Table 7-3, to some extent these results suggest that there is very little confounding of asbestos by chemicals other than the seven currently suspected lung carcinogens used in strategy 3. In contrast, some chemicals did appear to be confounded by other occupational chemicals, possibly highlighting the inadequacy of the simpler models that made more restrictive assumptions about which chemicals would be considered confounders.

For the most part, the addition of the categories of exchangeability (comparing estimates from strategy 7 to estimates from strategy 8) resulted in only minor changes in the estimates. Slightly larger influences can be seen in the results of a few chemicals, such as chromium fumes (104), whose estimate in strategy 7 (without the categories) was 1.4, but shifted upwards to 1.9 in strategy 8.

7.5.1 Estimates of the effects of chemical and physical properties on the risk of lung cancer

Visual inspection

Discerning patterns from a visual inspection of the estimates by categories in Table 7-11 is difficult because several of the categories involved a large number of substances. Nevertheless, in strategy 8, it appeared that there were some categories that predominantly contained chemicals with estimates that were 'elevated', though not necessarily statistically so, over the estimates from strategy 7. These categories included polypeptides, metal dusts (excluding oxides), metal oxide dusts, heavy metal compounds, aliphatic chlorinated hydrocarbons, organic gases (C1-C4), chromates, copper compounds, and zinc compounds.

Second-level semi-Bayes estimates

A more quantitative appreciation of the influence of these categories is possible by inspecting the estimates from the second-level model of strategy 8. These second-level coefficients, for the regression models of both levels of exposure, are provided in Table 7-12. The estimates are to be read as would be normal coefficients in a linear regression;

that is, expected change in the logistic coefficient for a particular chemical if that chemical belonged to that particular category, depending on the exposure level. This translates to the effects on the risk of lung cancer by the second-level covariates, the chemical and physical properties. For instance, the estimate for heavy metal compounds at any level of exposure was 0.214. Thus, exp(0.214)=1.2 is the expected rate ratio estimate for the effect of heavy metal compounds on the risk of lung cancer. With the exception of Z1, all the category-covariates were dichotomous: chemicals either belonged to the category or they did not. Z1, on the other hand, represented previous evidence (of lung carcinogenicity of a chemical) in the form of a continuous covariate (see step iii, section 6.10). Its associated parameter estimate has a curious interpretation, in that it is the expected change in the log-odds of lung cancer for a one unit increase in the variable, which would be the case if the prior belief for the effect of the chemical was a 'residual' relative risk of 2.7 (corresponding approximately to e^1).

Table 7-12: Estimated coefficients of second-level variables in semi-Bayes analyses

Variable	*Second-level characteristic, used to specify exchangeable effects*	*Parameter estimates*	
		Any exposure	*Substantial exposure*
Z1	Previous evidence	-0.460	0.437
Z2	Polypeptides	0.030	0.190
Z3	Polysaccharides	0.117	-0.038
Z4	Fibrous inorganic dusts	-0.144	-0.272
Z5	Silica containing compounds	0.184	0.111
Z6	Metal dusts (excluding oxides)	-0.007	0.081
Z7	Metal oxide dusts	0.061	0.388
Z8	Metal oxide fumes	-0.088	-0.020
Z9	Heavy metal compounds	0.214	-0.017
Z10	Monocyclic aromatic hydrocarbons	-0.207	-0.133
Z11	Polycyclic aromatic hydrocarbons	0.040	0.009
Z12	Engine emissions	0.050	0.056
Z13	Inorganic acid mists	-0.052	-0.171
Z14	Resins and resin-containing compounds	0.064	-0.036
Z15	Carbonaceous compounds	-0.212	-0.037
Z16	Aliphatic alkanes (C5-C17)	-0.004	0.318
Z17	Aliphatic alcohols	0.016	0.270
Z18	Aliphatic chlorinated hydrocarbons	0.084	0.432
Z19	Inorganic gases	0.024	0.034
Z20	Organic gases (C1-C4)	0.030	0.568
Z21	Inorganic salts	-0.006	0.254

Variable	*Second-level characteristic, used to specify exchangeable effects*	*Parameter estimates*	
		Any exposure	*Substantial exposure*
Z22	Magnesium compounds	-0.018	
Z23	Aluminium compounds	0.179	0.023
Z24	Chromates	0.719	
Z25	Manganese compounds	0.331	0.792
Z26	Iron compounds	-0.243	-0.618
Z27	Nickel compounds	-0.197	
Z28	Copper compounds	0.022	0.376
Z29	Zinc compounds	0.103	0.360
Z30	Tin compounds	-0.305	-0.712
Z31	Lead compounds	-0.087	-0.123

7.5.2 An example of an approximate semi-Bayes calculation

The estimates shown in Table 7-12 were used to shrink first-level chemical estimates to supposedly more accurate point values. Using lead chromate (40) as an example, the semi-Bayes estimate would be calculated by substituting lead chromate's z-matrix values into the estimated second-level regression equation and then averaging with the first-level maximum likelihood estimate. Lead chromate belonged to three categories of exchangeability: heavy metal compounds (Z9), chromates (Z24), and lead compounds (Z31), and in addition had a non-zero value for the intercept (Z1), which in this case encapsulated the effect of how I specified previous evidence of lung carcinogenicity (see step iii, section 6.10). Thus, at any level of exposure, the following equation would approximately represent the semi-Bayes estimate for lead chromate:

$$b_{sb,any} = [-0.460(Z1) + 0.214(Z9) + 0.719(Z24) - 0.087(Z31)]w_1 + 0.240w_2 ,$$

where the coefficients for Z1, Z9, Z24, and Z31 were reported in Table 7-12, 0.240 was the original maximum likelihood estimate (from model 6, results not shown), and the weights, w_1 and w_2, were inversely proportional, respectively, to the specified prior variance, T^2=0.246, and the estimated variance of the maximum likelihood estimate for lead chromate, in this case being 0.24. Values of 1 were substituted in for Z9, Z24, and Z31, indicating lead chromate belonged to those particular categories (the zeroes for the other Z-variables explain why only these four variables appear in the above equation). And the value of 0.8 was substituted into Z1, indicating that the prior belief for the effect of lead chromate was a rate ratio of $e^{0.8} \sim 2.2$ (see step iii, section 6.10). Substituting in all the values led to,

$b_{sb,any}$= [-0.460(0.8) +0.214(1) + 0.719(1)-0.087(1)]w_1 + 0.240w_2 ,

$b_{sb,any}$= 0.477w_1 + 0.240w_2 ,

Thus, depending on the weighting, the semi-Bayes estimate for lead chromate would fall somewhere between the rate ratios of exp(0.477) and exp(0.240). The semi-Bayes estimate for lead chromate, from Table 7-11, was log(1.367)=0.313, which falls almost halfway between the estimated prior mean and maximum likelihood estimate.

7.6 Ranking and selection of occupational substances

One of the objectives of the thesis was to earmark occupational substances that were most likely involved in the etiology of lung cancer and that would be prioritized for further examination. Although there are several ways of ranking and selecting, for the reasons outlined in section 6.11, the chemicals were primarily ranked according to the magnitude of the estimates yielded by the semi-Bayes approach of strategy 8, resulting in sections A and B of Table 7-13. Section A shows the selected chemicals from the results at any level of exposure, while section B shows the selected chemicals from the results at the substantial level of exposure. From a more conventional modeling approach to selecting the chemicals (strategy 3), sections C and D correspond, respectively, to the results from any level of exposure and the substantial level of exposure.

Special notice should be given to which chemicals were considered with each modeling strategy. The approach of modeling strategy 3 analyzed the entire set of 231 chemicals, while the semi-Bayes approach was restricted to the 184 chemicals that could be assessed simultaneously in a single regression model. For this reason, more chemicals could have been earmarked in sections C and D, which were the results from strategy 3.

Table 7-13: Four variations of the ranking and selection of occupational chemicals

A. Ranking by point estimate (STRATEGY 8, semi-Bayes regression), for ANY level of exposure, with selection by statistical significance, tail area values, and previous evidence.

	Reason for Selection [a]	Any Exposure			Substantial Exposure		
		Exp Cases	RR (90% CL)	AN[b]	Exp Cases	RR (90% CL)	AN[b]
53. Natural Rubber	Sig	44	**2.1 (1.3, 3.5)**	23.1		NE [c]	
83. Nitrogen Oxides	Sig	240	**2.1 (1.5, 2.9)**	123.4	36	1.6 (1.0, 2.8)	14.1
104. Chromium Fumes	Tail	43	**1.9 (0.6, 6.1)**	20.3	8	1.6 (0.7, 4.0)	3.1
12. Clay Dust	Sig	28	**1.8 (1.0, 3.2)**	12.5	9	1.2 (0.6, 2.7)	1.7
23. Cosmetic Talc	Tail	13	**1.6 (0.8, 3.2)**	5.1		NE [c]	
68. Urea-Formald.	Tail	50	**1.6 (1.0, 2.5)**	18.6	4	1.0 (0.5, 2.0)	0
64. Poly-Acrylates	Tail	29	**1.6 (0.9, 2.8)**	10.5	8	1.3 (0.6, 2.5)	1.6
108. Copper Fumes	Tail	47	**1.5 (0.9, 2.7)**	16.3	16	1.6 (0.7, 3.6)	5.8
127. Alkali, Caustic Solutions	Sig	72	**1.5 (1.1, 2.2)**	24.7	23	1.6 (1.0, 2.6)	8.4
40. Lead Chromate	Prev	35	**1.4 (0.7, 2.8)**	9.4		NE [c]	
117. Diesel Eng.Emissions	Prev	165	**1.1 (0.9, 1.4)**	17.4	81	1.2 (0.8, 1.7)	11.9
219. Benzo(a)pyrene	Prev	220	**1.1 (0.8, 1.5)**	19.1	42	1.0 (0.5, 1.7)	0
6. Crystalline Silica	Prev	238	**1.0 (0.8, 1.3)**	2.3	81	1.3 (0.9, 1.9)	17.0
195. Cadmium Compounds	Prev	11	**1.0 (0.5, 2.2)**	0.1	4	1.1 (0.4, 2.8)	0.3

[a] Selection of which chemicals to include in this table depended on the analytic approach. Selection of estimates required a RR>1.0 and one of the following: statistically significant estimate with P-value≤0.1 (Sig), previous evidence as to what are currently thought of as lung carcinogens (Prev), or -- only in the case of the semi-Bayes models -- a point estimate in the extreme tail region of the distribution of estimates (Tail).
[b] Attributable number. When RR ≤ 1, AN=0. [c] Not estimated at this exposure level.

Table 7-13: Four variations of the ranking and selection of occupational chemicals

B. Ranking by point estimate (STRATEGY 8, semi-Bayes regression), for the SUBSTANTIAL level of exposure, with selection by statistical significance, tail area values, and previous evidence.

		Any Exposure			Substantial Exposure		
	Reason for Selection [a]	Exp Cases	RR (90% CL)	AN[b]	Exp Cases	RR (90% CL)	AN[b]
105. Manganese Fumes	Tail	60	1.2 (0.7, 2.2)	10.5	14	**2.2 (0.9, 5.5)**	7.7
91. Methane	Tail	41	1.0 (0.6, 1.7)	1.5	13	**2.1 (0.9, 4.9)**	6.8
15. Brass Dust	Tail	24	1.3 (0.8, 2.3)	6.1	10	**2.0 (0.9, 4.5)**	5.0
35. Copper Dust	Tail	47	1.2 (0.8, 2.0)	9.0	14	**1.9 (0.9, 3.9)**	6.6
156. Kerosene	Sig	69	1.2 (0.9, 1.7)	12.7	26	**1.8 (1.1, 3.1)**	11.8
104. Chromium Fumes	Prev	43	1.9 (0.6, 6.1)	20.3	8	**1.6 (0.7, 4.0)**	3.1
5. Asbestos	Prev	177	1.0 (0.7, 1.3)	0	34	**1.4 (0.9, 2.4)**	10.3
6. Crystalline Silica	Prev	238	1.0 (0.8, 1.3)	2.3	81	**1.3 (0.9, 1.9)**	17.0
214. PAH (Any)	Prev	581	1.0 (0.7, 1.3)	0	80	**1.2 (0.8, 1.9)**	15.6
193. Arsenic Compounds	Prev	31	0.7 (0.4, 1.3)	0	11	**1.2 (0.5, 2.7)**	1.9
117. Diesel Eng.Emissions	Prev	165	1.1 (0.9, 1.4)	17.4	81	**1.2 (0.8, 1.7)**	11.9
195. Cadmium Compounds	Prev	11	1.0 (0.5, 2.2)	0.1	4	**1.1 (0.4, 2.8)**	0.3
163. Coal Tar and Pitch	Prev	23	0.9 (0.6, 1.5)	0	12	**1.1 (0.6, 2.0)**	0.7

[a] Selection of which chemicals to include in this table depended on the analytic approach. Selection of estimates required a RR>1.0 and one of the following: statistically significant estimate with P-value≤0.1 (Sig), previous evidence as to what are currently thought of as lung carcinogens (Prev), or -- only in the case of the semi-Bayes models -- a point estimate in the extreme tail region of the distribution of estimates (Tail).
[b] Attributable number. When RR ≤ 1, AN=0. [c] Not estimated at this exposure level.

Table 7-13: Four variations of the ranking and selection of occupational chemicals

C. Ranking by point estimate (STRATEGY 3, conventional logistic regression), for ANY level of exposure, with selection by statistical significance and previous evidence.

		Any Exposure			Substantial Exposure		
	Reason for Selection [a]	Exp Cases	RR (90% CL)	AN[b]	Exp Cases	RR (90% CL)	AN[b]
104. Chromium Fumes	Sig	43	**2.0 (1.2, 3.2)**	21.0	8	2.1 (0.7, 6.6)	4.1
182. Magnesium Compounds	Sig	19	**1.9 (1.1, 3.3)**	8.8	6	1.9 (0.7, 4.9)	2.8
8. Glass Dust	Sig	18	**1.9 (1.0, 3.3)**	8.3	6	0.7 (0.3, 1.8)	0
108. Copper Fumes	Sig	47	**1.8 (1.2, 2.7)**	21.2	16	2.1 (1.1, 3.9)	8.3
107. Nickel Fumes	Sig	42	**1.7 (1.1, 2.8)**	17.7		NE [c]	
171. Cutting Fluids post 1955	Sig	62	**1.7 (1.2, 2.4)**	25.8	25	1.4 (0.9, 2.3)	7.1
85. Hydrogen Fluoride	Sig	38	**1.7 (1.1, 2.6)**	15.8	4	1.3 (0.4, 4.3)	0.9
177. Fluorides	Sig	42	**1.6 (1.1, 2.4)**	16.1	6	1.4 (0.5, 3.8)	1.8
83. Nitrogen Oxides	Sig	240	**1.6 (1.3, 2.0)**	91.4	36	1.6 (1.0, 2.5)	13.8
122. Propane Eng.Emiss.	Sig	28	**1.6 (1.0, 2.5)**	10.5	15	1.1 (0.6, 2.0)	1.7
156. Kerosene	Sig	69	**1.6 (1.2, 2.1)**	25.5	26	2.6 (1.5, 4.4)	15.8
174. Inks	Sig	37	**1.5 (1.1, 2.3)**	13.1	17	1.1 (0.7, 1.9)	2.1
195. Cadmium Compounds	Prev	11	**1.5 (0.7, 3.0)**	3.6	4	1.4 (0.5, 4.4)	1.2
190. Nickel Compounds	Sig	79	**1.5 (1.1, 2.0)**	25.3	12	1.7 (0.7, 3.8)	4.9
21. Aluminium Alloy Dust	Sig	63	**1.4 (1.0, 1.9)**	17.4	25	1.1 (0.7, 1.8)	2.7
204. Alkanes (C5-C17)	Sig	368	**1.4 (1.1, 1.6)**	99.2	207	1.5 (1.2, 1.8)	67.4
98. Gas Welding Fumes	Sig	115	**1.3 (1.0, 1.7)**	29.4	50	1.4 (1.0, 2.0)	13.4
178. Chromium (VI) Comp.	Sig	90	**1.3 (1.0, 1.7)**	21.9	12	1.5 (0.8, 2.8)	3.8
161. Cutting Fluids	Sig	85	**1.3 (1.0, 1.7)**	20.5	33	1.2 (0.8, 1.7)	4.5
196. Tin Compounds	Sig	92	**1.3 (1.0, 1.7)**	21.4	14	1.1 (0.6, 2.0)	1.3

[a] Selection of which chemicals to include in this table depended on the analytic approach. Selection of estimates required a RR>1.0 and one of the following: statistically significant estimate with P-value≤0.1 (Sig), previous evidence as to what are currently thought of as lung carcinogens (Prev), or -- only in the case of the semi-Bayes models -- a point estimate in the extreme tail region of the distribution of estimates (Tail).
[b] Attributable number. When RR ≤ 1, AN=0. [c] Not estimated at this exposure level.

Table 7-13: Four variations of the ranking and selection of occupational chemicals

183. Aluminium Compounds	Sig	199	**1.3 (1.0, 1.6)**	43.4	51	1.4 (1.0, 2.0)	14.3
4. Metallic Dust	Sig	276	**1.3 (1.1, 1.5)**	57.7	138	1.4 (1.1, 1.8)	38.5
6. Crystalline Silica	Sig	238	**1.2 (1.0, 1.5)**	42.6	81	1.6 (1.2, 2.2)	30.9
117. Diesel Eng.Emissions	Prev	165	**1.1 (0.9, 1.4)**	20.3	81	1.1 (0.9, 1.5)	10.1
219. Benzo(a)pyrene	Prev	220	**1.1 (0.9, 1.4)**	24.4	42	1.0 (0.6, 1.6)	0
5. Asbestos	Prev	177	**1.0 (0.8, 1.3)**	7.1	34	1.5 (1.0, 2.4)	11.7
214. PAH (Any)	Prev	581	**1.0 (0.9, 1.3)**	22.7	80	1.2 (0.8, 1.8)	15.2
76. Soot	Prev	91	**1.0 (0.7, 1.4)**	0.2	26	1.6 (0.9, 2.7)	9.3

[a] Selection of which chemicals to include in this table depended on the analytic approach. Selection of estimates required a RR>1.0 and one of the following: statistically significant estimate with P-value≤0.1 (Sig), previous evidence as to what are currently thought of as lung carcinogens (Prev), or -- only in the case of the semi-Bayes models -- a point estimate in the extreme tail region of the distribution of estimates (Tail).
[b] Attributable number. When RR ≤ 1, AN=0. [c] Not estimated at this exposure level.

Table 7-13: Four variations of the ranking and selection of occupational chemicals

D. Ranking by point estimate (STRATEGY 3, conventional logistic regression), for the SUBSTANTIAL level of exposure, with selection by statistical significance and previous evidence.

		Any Exposure			Substantial Exposure		
	Reason for Selection [a]	Exp Cases	RR (90% CL)	AN[b]	Exp Cases	RR (90% CL)	AN[b]
24. Borates	Sig	11	1.7 (0.9, 3.5)	4.6	4	**4.1 (1.1, 15.9)**	3.0
36. Zinc Dust	Sig	26	1.4 (0.9, 2.3)	8.0	5	**3.5 (1.1, 11.1)**	3.6
213. Glycol Ethers	Sig	26	1.0 (0.6, 1.5)	0	8	**3.3 (1.2, 8.9)**	5.6
105. Manganese Fumes	Sig	60	1.4 (1.0, 1.9)	16.6	14	**3.3 (1.5, 7.0)**	9.7
90. Natural Gas	Sig	24	0.9 (0.6, 1.4)	0	8	**3.1 (1.1, 8.3)**	5.4
187. Manganese Compounds	Sig	71	1.3 (0.9, 1.7)	15.1	14	**2.9 (1.4, 6.0)**	9.1
35. Copper Dust	Sig	47	1.2 (0.8, 1.7)	6.8	14	**2.6 (1.3, 5.3)**	8.7
91. Methane	Sig	41	0.9 (0.6, 1.3)	0	13	**2.6 (1.3, 5.4)**	8.0
109. Zinc Fumes	Sig	39	1.3 (0.9, 1.9)	8.9	16	**2.6 (1.3, 5.1)**	9.8
156. Kerosene	Sig	69	1.6 (1.2, 2.1)	25.5	26	**2.6 (1.5, 4.4)**	15.8
192. Zinc Compounds	Sig	107	1.2 (0.9, 1.5)	17.3	25	**2.1 (1.2, 3.6)**	13.1
108. Copper Fumes	Sig	47	1.8 (1.2, 2.7)	21.2	16	**2.1 (1.1, 3.9)**	8.3
104. Chromium Fumes	Prev	43	2.0 (1.2, 3.2)	21.0	8	**2.1 (0.7, 6.6)**	4.1
158. Heating Oil	Sig	53	1.3 (0.9, 1.9)	13.0	24	**1.9 (1.1, 3.3)**	11.7
83. Nitrogen Oxides	Sig	240	1.6 (1.3, 2.0)	91.4	36	**1.6 (1.0, 2.5)**	13.8
6. Crystalline Silica	Sig	238	1.2 (1.0, 1.5)	42.6	81	**1.6 (1.2, 2.2)**	30.9
191. Copper Compounds	Sig	128	1.3 (1.0, 1.6)	26.5	55	**1.6 (1.1, 2.3)**	20.2
76. Soot	Prev	91	1.0 (0.7, 1.4)	0.2	26	**1.6 (0.9, 2.7)**	9.3
5. Asbestos	Prev	177	1.0 (0.8, 1.3)	7.1	34	**1.5 (1.0, 2.4)**	11.7
3. Excavation Dust	Sig	109	1.3 (1.0, 1.8)	27.3	69	**1.5 (1.1, 2.2)**	23.6

[a] Selection of which chemicals to include in this table depended on the analytic approach. Selection of estimates required a RR>1.0 and one of the following: statistically significant estimate with P-value≤0.1 (Sig), previous evidence as to what are currently thought of as lung carcinogens (Prev), or -- only in the case of the semi-Bayes models -- a point estimate in the extreme tail region of the distribution of estimates (Tail).
[b] Attributable number. When RR ≤ 1, AN=0. [c] Not estimated at this exposure level.

Table 7-13: Four variations of the ranking and selection of occupational chemicals

204. Alkanes (C5-C17)	Sig	368	1.4 (1.1, 1.6)	99.2	207	**1.5 (1.2, 1.8)**	67.4
101. Metal Oxide Fumes	Sig	190	1.2 (1.0, 1.4)	28.3	87	**1.4 (1.1, 1.9)**	26.7
195. Cadmium Compounds	Prev	11	1.5 (0.7, 3.0)	3.6	4	**1.4 (0.5, 4.4)**	1.2
97. Coal Gas	Prev	8	0.6 (0.3, 1.1)	0	2	**1.4 (0.3, 6.8)**	0.6
4. Metallic Dust	Sig	276	1.3 (1.1, 1.5)	57.7	138	**1.4 (1.1, 1.8)**	38.5
214. PAH (Any)	Prev	581	1.0 (0.9, 1.3)	22.7	80	**1.2 (0.8, 1.8)**	15.2
117. Diesel Eng.Emissions	Prev	165	1.1 (0.9, 1.4)	20.3	81	**1.1 (0.9, 1.5)**	10.1
40. Lead Chromate	Prev	35	0.7 (0.4, 1.2)	0	3	**1.1 (0.2, 5.6)**	0.2
163. Coal Tar and Pitch	Prev	23	0.9 (0.5, 1.4)	0	12	**1.0 (0.5, 2.3)**	0.4

[a] Selection of which chemicals to include in this table depended on the analytic approach. Selection of estimates required a RR>1.0 and one of the following: statistically significant estimate with P-value≤0.1 (Sig), previous evidence as to what are currently thought of as lung carcinogens (Prev), or -- only in the case of the semi-Bayes models -- a point estimate in the extreme tail region of the distribution of estimates (Tail).
[b] Attributable number. When RR ≤ 1, AN=0. [c] Not estimated at this exposure level.

The four different approaches in Table 7-13 provided four different selections of chemicals. Aside from a few chemicals being chosen in all cases based on previous evidence, there were no other chemicals that appeared on both of the semi-Bayes lists; the chemicals earmarked at any level of exposure were different from those earmarked at the substantial level of exposure. Even the currently suspected carcinogens were not, in their entirety, reproduced on both lists. This is because many of the chemicals that are currently suspected of being lung carcinogens had, in the analyses, point estimates that were just below unity, precluding them from appearing on these lists.

The conventional approach (strategy 3) produced many more statistically significant estimates than the semi-Bayes approach. One explanation is the fact that 231 chemicals were considered in strategy 3, over the 184 chemicals in strategy 8. For example, magnesium compounds (182) is second on the list of section C, but was not included at all in the semi-Bayes modeling, and so could not appear on the lists of sections A and B.

As expected, the estimates from strategy 3 did not have the same distributional properties as the estimates from model 8. Many more statistically significant estimates occurred, but these estimates did not predominantly occur in the extreme tails of the distributions as they did in the semi-Bayes models. As a result, had the 'tail region' criterion been implemented to select estimates from strategy 3 (see section 6.11.2), more than twice as many chemicals would have appeared as there currently are in sections C and D.

Comparing the results of parts C and D, based on the selection of estimates from the two exposure levels of modeling strategy 3, several chemicals showed consistently elevated estimates with P-value ≤ 0.1. These included chromium fumes (104), copper fumes (108), nitrogen oxides (83), kerosene (156), alkanes (C5-C17) (204), and metallic dust (4).

Comparing across the conventional and semi-Bayes strategies, nitrogen oxides (83) and copper fumes (108) were both earmarked at any level of exposure, while manganese fumes (105) and copper dust (35) were both earmarked at the substantial level of exposure.

An estimate of the numbers of cases attributable to each of the chemical exposures was provided as a means of appreciating the impact of exposure in the study population. As

an alternative, ranking by the attributable number (AN) would have produced a very different picture as to which chemicals would have been selected. Noteworthy are the results for nitrogen oxides and alkanes (C5-C17), both of which were highly prevalent in the study population and had high estimates for AN across most of the modeling strategies.

The estimated values for the AN were not always consistent with logic. For example, in section A of the semi-Bayes results, crystalline silica (6) had an AN at any level of exposure of 2.3, while at the substantial level of exposure it was 17.0. As the latter exposure level was subsumed by the former, this result was not logically possible, highlighting the instability of such empirical estimates.

Combining the four lists resulted in a total of 53 chemicals, which were carried forward for further analyses. Of the 53, 11 were included solely because they are currently recognized as or suspected of being lung carcinogens, while the remaining 42 chemicals had supporting evidence from the present results.

7.7 Secondary analyses on selected chemicals

The following analyses were only applied to the 53 chemicals earmarked in the previous section. In some cases, results for particular chemicals were omitted if the estimate was deemed too imprecise and unstable, which tended to occur among those chemicals with low prevalence in the study population.

7.7.1 Concentration, duration, and time windows of exposure

Results presented in the previous sections all relied on a composite index of exposure, which combined the exposure characteristics of concentration, frequency, duration, and certainty into a cumulative lifetime index. To explore the different characteristics of exposure, Table 7-14 presents estimates of the rate ratios by concentration, duration, and two time windows of exposure (see section 6.13.1).

Many of the estimates in the table were imprecise and somewhat unstable due to the more finely restricted definitions of variables for these analyses. Nevertheless, substances whose estimates suggested a response curve with increasing concentration, included: metallic dust, copper dust, zinc dust, soot, natural gas, gas welding fumes, metal oxide

fumes, zinc fumes, kerosene, magnesium compounds, zinc compounds, glycol ethers, and any source of polycyclic aromatic hydrocarbons (PAH).

In a similar fashion, two categories of duration of exposure were assessed. Only a few substances showed a clear pattern of stronger associations with lung cancer when the cumulative duration of exposure exceeded ten years: excavation dust, crystalline silica, heating oil, and manganese compounds.

The effects of exposures in two different windows of time were also analyzed. As in the previous analyses, the data was occasionally too sparse to support analyses of separate windows of time. Thus, too few people were exposed to borates, for example, in the window proximal to diagnosis, so the resulting estimate was too unstable to report. Although an estimate for borates was shown for the exposure window of twenty or more years prior to diagnosis, even that estimate might be considered too imprecise to be meaningful. Some of these results are nevertheless suggestive of certain periods being more relevant for the effects of the exposure. Substances showing their strongest association with lung cancer for exposure in the 5 to 15 years before diagnosis, included: excavation dust, metallic dust, asbestos, glass dust, urea-formaldehyde, hydrogen fluoride, chromium fumes, manganese fumes, nickel fumes, copper fumes, kerosene, cutting fluids, cutting fluids post 1955, fluorides, aluminium compounds, manganese compounds, nickel compounds, and cadmium compounds. Substances showing their strongest association with lung cancer for exposures more than 20 years before diagnosis, included: clay dust, brass dust, and aluminium alloy dust.

Table 7-14: Rate ratio estimates, for selected chemicals, by concentration, duration, and windows of exposure, with adjustment for non-occupational confounders

		Concentration [a]		Duration		Time windows prior to diagnosis	
		Low/medium	High	1-10 years	10+ years	5-20 years	20+ years
		RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)
3.	Excavation Dust	1.3 (0.8, 2.1)	1.5 (1.1, 1.9)	1.1 (0.8, 1.6)	1.8 (1.3, 2.5)	1.8 (1.0, 3.2)	1.1 (0.8, 1.6)
4.	Metallic Dust	1.2 (1.0, 1.5)	1.5 (1.2, 1.8)	1.3 (1.0, 1.6)	1.4 (1.1, 1.7)	1.5 (1.0, 2.2)	1.2 (0.9, 1.6)
5.	Asbestos	1.2 (0.9, 1.4)	1.2 (0.8, 1.9)	1.2 (0.9, 1.7)	1.1 (0.9, 1.4)	2.0 (1.2, 3.3)	1.0 (0.7, 1.5)
6.	Crystalline Silica	1.2 (1.0, 1.4)	1.4 (1.1, 1.9)	1.0 (0.8, 1.3)	1.4 (1.1, 1.7)	1.2 (0.8, 2.0)	1.0 (0.8, 1.3)
8.	Glass Dust	4.2 (1.8, 10.3)	1.1 (0.5, 2.2)	1.8 (0.8, 4.2)	1.9 (0.9, 4.2)	4.1 (1.1, 15.0)	1.4 (0.5, 3.8)
12.	Clay Dust	2.6 (1.2, 5.4)	1.6 (0.9, 2.8)	1.6 (0.9, 3.0)	2.4 (1.2, 4.7)	1.2 (0.4, 3.8)	2.0 (1.0, 3.8)
15.	Brass Dust	1.3 (0.7, 2.7)	1.8 (0.9, 3.3)	2.5 (1.1, 5.3)	1.2 (0.7, 2.2)	0.6 (0.1, 3.9)	2.4 (1.1, 5.1)
21.	Aluminium Alloy Dust	1.6 (1.0, 2.4)	1.3 (0.9, 2.0)	1.6 (1.1, 2.5)	1.3 (0.9, 1.9)	1.0 (0.5, 2.1)	1.6 (1.0, 2.5)
23.	Cosmetic Talc	2.3 (1.1, 4.9)	0.4 (0.1, 1.4)	3.6 (0.9, 13.7)	1.0 (0.5, 2.0)	12.0 (1.6, 89.6)	4.0 (1.1, 15.0)
24.	Borates	1.6 (0.7, 3.6)	3.2 (0.9, 11.3)	1.0 (0.3, 4.4)	2.4 (1.1, 5.2)		0.9 (0.2, 3.5)
35.	Copper Dust	1.0 (0.6, 1.4)	2.5 (1.3, 4.7)	1.1 (0.7, 1.8)	1.3 (0.8, 2.1)	0.9 (0.4, 2.0)	1.2 (0.7, 2.1)
36.	Zinc Dust	1.2 (0.7, 2.1)	5.8 (1.9, 17.7)	2.4 (1.1, 5.0)	1.3 (0.8, 2.3)	1.3 (0.3, 6.0)	1.8 (0.9, 3.6)
40.	Lead Chromate	1.1 (0.7, 1.6)	1.5 (0.4, 5.3)	1.2 (0.6, 2.3)	1.1 (0.7, 1.7)	1.2 (0.5, 2.9)	1.1 (0.5, 2.3)
53.	Natural Rubber	1.5 (1.0, 2.2)	0.5 (0.2, 1.1)	1.6 (0.9, 2.6)	0.9 (0.6, 1.5)	1.2 (0.4, 3.6)	1.2 (0.7, 2.2)
64.	Poly-Acrylates	1.3 (0.8, 2.1)	1.1 (0.5, 2.4)	1.2 (0.5, 2.8)	1.3 (0.8, 2.0)	1.2 (0.5, 2.5)	0.5 (0.1, 3.6)

All estimates are in reference to 'never exposed'

[a] High concentration was defined by 3-point chemist codings for concentration and frequency as both above 2 (high); low/medium concentration was defined as any combination of the 3-point chemist codings for concentration and frequency if at least one was between 1 and 2 (low to medium).

Table 7-14: Rate ratio estimates, for selected chemicals, by concentration, duration, and windows of exposure, with adjustment for non-occupational confounders

	Concentration [a]		Duration		Time windows prior to diagnosis	
	Low/medium	High	1-10 years	10+ years	5-20 years	20+ years
	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)
68. Urea-Formald.	1.5 (1.1, 2.2)	1.0 (0.4, 2.5)	1.5 (0.8, 2.7)	1.4 (1.0, 2.1)	2.2 (1.0, 5.2)	1.3 (0.6, 3.0)
76. Soot	1.0 (0.7, 1.3)	1.8 (1.2, 2.8)	1.3 (0.9, 1.9)	1.1 (0.8, 1.5)	1.4 (0.8, 2.8)	1.0 (0.7, 1.5)
83. Nitrogen Oxides	1.6 (1.3, 1.9)	1.4 (0.9, 2.1)	1.9 (1.4, 2.4)	1.4 (1.2, 1.8)	2.4 (1.6, 3.7)	1.5 (1.2, 2.0)
85. Hydrogen Fluoride	2.2 (1.4, 3.4)	1.2 (0.5, 3.4)	1.8 (0.9, 3.7)	2.1 (1.3, 3.4)	3.8 (1.5, 9.6)	1.1 (0.5, 2.4)
90. Natural Gas	0.8 (0.5, 1.2)	4.2 (1.4, 12.5)	2.2 (1.0, 4.6)	0.7 (0.4, 1.3)	1.3 (0.6, 2.9)	1.2 (0.2, 9.4)
91. Methane	0.9 (0.6, 1.4)	1.5 (0.8, 3.0)	1.2 (0.7, 2.0)	0.9 (0.6, 1.4)	1.4 (0.7, 3.1)	0.8 (0.4, 1.5)
97. Coal Gas	0.4 (0.2, 1.1)	1.4 (0.4, 4.8)	0.7 (0.2, 2.0)	0.5 (0.2, 1.4)		0.4 (0.1, 1.3)
98. Gas Welding Fumes	1.3 (0.9, 1.7)	1.8 (1.3, 2.4)	1.6 (1.1, 2.4)	1.4 (1.1, 1.8)	1.8 (1.1, 3.0)	1.7 (1.1, 2.6)
101. Metal Oxide Fumes	1.0 (0.8, 1.3)	1.7 (1.3, 2.2)	1.3 (1.0, 1.7)	1.3 (1.0, 1.6)	1.4 (0.9, 2.1)	1.2 (0.9, 1.7)
104. Chromium Fumes	2.3 (1.5, 3.6)	2.0 (0.9, 4.4)	2.4 (1.2, 4.9)	2.2 (1.4, 3.5)	4.4 (1.8, 10.7)	1.9 (0.9, 4.1)
105. Manganese Fumes	1.5 (1.1, 2.2)	2.0 (1.1, 3.7)	1.3 (0.8, 2.1)	1.9 (1.3, 2.9)	2.4 (1.1, 4.9)	1.1 (0.7, 1.9)
107. Nickel Fumes	2.3 (1.5, 3.6)	1.5 (0.7, 3.4)	1.9 (0.9, 3.8)	2.3 (1.4, 3.5)	3.8 (1.6, 8.8)	1.5 (0.7, 3.4)
108. Copper Fumes	2.0 (1.2, 3.3)	2.3 (1.3, 4.0)	2.4 (1.3, 4.4)	2.0 (1.3, 3.2)	2.9 (1.1, 7.7)	1.6 (0.8, 3.0)
109. Zinc Fumes	1.2 (0.7, 1.9)	2.7 (1.5, 4.8)	1.2 (0.7, 2.3)	1.9 (1.2, 3.0)	1.7 (0.7, 4.1)	1.0 (0.5, 1.9)
117. Diesel Eng.Emissions	1.2 (0.9, 1.5)	1.2 (0.9, 1.5)	1.3 (1.0, 1.9)	1.1 (0.9, 1.4)	1.2 (0.8, 1.8)	1.4 (0.9, 2.1)

All estimates are in reference to 'never exposed'

[a] High concentration was defined by 3-point chemist codings for concentration and frequency as both above 2 (high); low/medium concentration was defined as any combination of the 3-point chemist codings for concentration and frequency if at least one was between 1 and 2 (low to medium).

Table 7-14: Rate ratio estimates, for selected chemicals, by concentration, duration, and windows of exposure, with adjustment for non-occupational confounders

	Concentration [a]		Duration		Time windows prior to diagnosis	
	Low/medium	High	1-10 years	10+ years	5-20 years	20+ years
	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)
122. Propane Eng.Emiss.	4.4 (2.0, 9.5)	1.2 (0.7, 2.0)	1.8 (0.8, 4.0)	1.7 (1.0, 2.9)	1.9 (0.9, 4.0)	1.7 (0.1, 24.6)
127. Alkali, Caustic Solutions	1.4 (1.0, 2.0)	1.2 (0.8, 1.7)	1.0 (0.7, 1.5)	1.6 (1.1, 2.3)	1.2 (0.6, 2.4)	0.9 (0.6, 1.5)
156. Kerosene	1.2 (0.8, 1.7)	2.6 (1.6, 4.2)	1.4 (0.9, 2.1)	1.8 (1.2, 2.7)	4.4 (1.2, 16.7)	1.3 (0.9, 1.8)
158. Heating Oil	1.2 (0.8, 1.9)	1.7 (1.1, 2.7)	1.0 (0.6, 1.6)	1.8 (1.2, 2.8)	1.7 (0.8, 3.6)	1.0 (0.6, 1.8)
161. Cutting Fluids	1.3 (0.9, 1.8)	1.4 (1.0, 1.9)	1.1 (0.7, 1.6)	1.5 (1.1, 2.1)	3.2 (1.6, 6.5)	0.9 (0.6, 1.3)
163. Coal Tar and Pitch	0.6 (0.2, 1.5)	1.3 (0.8, 2.3)	1.4 (0.7, 2.9)	0.9 (0.5, 1.6)	3.1 (0.4, 23.8)	1.3 (0.7, 2.6)
171. Cutting Fluids post 1955	1.7 (1.2, 2.6)	1.4 (0.9, 2.3)	1.3 (0.7, 2.5)	1.7 (1.2, 2.4)	3.1 (1.6, 6.2)	0.8 (0.3, 2.3)
174. Inks	2.5 (1.4, 4.5)	1.1 (0.7, 1.8)	1.5 (0.8, 2.7)	1.6 (1.0, 2.5)	1.8 (0.8, 4.2)	1.2 (0.6, 2.4)
177. Fluorides	2.1 (1.4, 3.1)	1.2 (0.5, 2.9)	1.6 (0.8, 3.0)	2.0 (1.3, 3.2)	3.4 (1.5, 7.7)	1.1 (0.6, 2.3)
178. Chromium (VI) Comp.	1.4 (1.1, 1.9)	1.5 (0.9, 2.6)	1.4 (1.0, 2.2)	1.4 (1.0, 1.9)	1.6 (0.9, 2.8)	1.3 (0.8, 2.0)
182. Magnesium Compounds	1.3 (0.6, 2.6)	3.7 (1.6, 8.8)	2.1 (0.8, 5.4)	1.8 (0.9, 3.7)	1.1 (0.1, 7.6)	1.3 (0.5, 3.1)
183. Aluminium Compounds	1.4 (1.1, 1.7)	1.6 (1.2, 2.2)	1.6 (1.2, 2.1)	1.3 (1.1, 1.7)	1.9 (1.3, 3.0)	1.1 (0.8, 1.5)
187. Manganese Compounds	1.4 (1.0, 1.9)	2.0 (1.1, 3.6)	1.2 (0.7, 1.8)	1.8 (1.2, 2.5)	2.5 (1.3, 4.9)	0.9 (0.5, 1.4)
190. Nickel Compounds	1.8 (1.3, 2.4)	1.3 (0.7, 2.5)	1.3 (0.8, 2.2)	1.9 (1.3, 2.5)	1.9 (1.0, 3.6)	1.2 (0.7, 2.2)
191. Copper Compounds	1.2 (0.9, 1.6)	1.5 (1.1, 2.0)	1.3 (0.9, 1.7)	1.3 (1.0, 1.7)	1.7 (0.9, 3.4)	0.9 (0.7, 1.3)

All estimates are in reference to 'never exposed'

[a] High concentration was defined by 3-point chemist codings for concentration and frequency as both above 2 (high); low/medium concentration was defined as any combination of the 3-point chemist codings for concentration and frequency if at least one was between 1 and 2 (low to medium).

Table 7-14: Rate ratio estimates, for selected chemicals, by concentration, duration, and windows of exposure, with adjustment for non-occupational confounders

	Concentration [a]		Duration		Time windows prior to diagnosis	
	Low/medium	High	1-10 years	10+ years	5-20 years	20+ years
	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)	RR (90% CL)
192. Zinc Compounds	1.2 (1.0, 1.6)	2.0 (1.2, 3.2)	1.5 (1.0, 2.2)	1.3 (1.0, 1.7)	1.0 (0.5, 2.0)	1.2 (0.8, 1.7)
193. Arsenic Compounds	0.7 (0.4, 1.1)	1.0 (0.6, 1.8)	0.9 (0.5, 1.5)	0.7 (0.4, 1.2)	0.5 (0.1, 3.5)	0.6 (0.4, 1.0)
195. Cadmium Compounds	2.0 (0.8, 5.1)	1.5 (0.6, 4.1)	1.5 (0.5, 4.5)	1.9 (0.8, 4.6)	9.8 (2.0, 48.8)	0.6 (0.2, 2.3)
196. Tin Compounds	1.4 (1.1, 1.9)	1.4 (0.8, 2.5)	1.6 (1.0, 2.3)	1.4 (1.0, 1.8)	2.4 (1.2, 4.9)	1.1 (0.7, 1.7)
204. Alkanes (C5-C17)	1.4 (1.1, 1.6)	1.5 (1.2, 1.8)	1.4 (1.1, 1.8)	1.4 (1.2, 1.7)	1.3 (0.9, 1.9)	1.3 (1.0, 1.7)
213. Glycol Ethers	0.9 (0.6, 1.5)	3.6 (1.4, 9.5)	1.7 (0.9, 3.2)	0.9 (0.5, 1.6)	0.8 (0.3, 1.9)	1.6 (0.7, 3.7)
214. PAH (Any)	1.1 (1.0, 1.3)	1.6 (1.1, 2.5)	1.1 (0.9, 1.5)	1.1 (1.0, 1.4)	0.8 (0.6, 1.2)	1.0 (0.8, 1.4)
219. Benzo(a)pyrene	1.2 (1.0, 1.4)	1.4 (1.0, 1.9)	1.2 (0.9, 1.6)	1.2 (0.9, 1.4)	1.4 (0.9, 2.4)	1.1 (0.9, 1.4)

All estimates are in reference to 'never exposed'

[a] High concentration was defined by 3-point chemist codings for concentration and frequency as both above 2 (high); low/medium concentration was defined as any combination of the 3-point chemist codings for concentration and frequency if at least one was between 1 and 2 (low to medium).

7.7.2 Occupational chemicals and histological subtypes of lung cancer

Of the 857 cases of lung cancer, 159 were diagnosed as small cell carcinomas, 359 as squamous cell carcinomas, and 167 as adenocarcinomas. The rest were diagnosed as other cell types, or the histological diagnosis was uncertain. The results of Table 7-15 show the subtype-specific estimates for the effects of the chemicals (see section 6.13.2), each only adjusted for the standard eight non-occupational confounders (modeling strategy 2).

Of the selected chemicals, many appeared to show results that would suggest certain chemicals were more related to one histological subtype of lung cancer over the others. Chemicals that showed evidence of being related predominantly to small cell carcinomas, included: excavation dust, asbestos, and alkali caustic solutions. Chemicals showing evidence of an association with squamous cell carcinomas, included: glass dust, brass dust, aluminium alloy dust, urea-formaldehyde, propane engine emissions, magnesium compounds, any source of polycyclic aromatic hydrocarbons, and benzo(a)pyrene. Finally, only borates showed evidence of an association primarily with adenocarcinomas.

Special notice should be given to the fact that these analyses focused on any level of exposure history, and did not assess the substantial exposure level. And even at any level of exposure, many of the estimates were too imprecise to report.

Table 7-15: Rate ratio estimates, from strategy 2, for selected chemicals, with respect to histological subtypes of lung cancer

	All lung cancer		Small cell		Squamous cell		Adenocarcinoma	
	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)
3. Excavation Dust	109	1.5 (1.2, 1.8)	29	2.5 (1.7, 3.8)	42	1.3 (0.9, 1.8)	11	0.7 (0.4, 1.2)
4. Metallic Dust	276	1.3 (1.1, 1.6)	64	2.0 (1.5, 2.7)	111	1.2 (1.0, 1.5)	54	1.3 (1.0, 1.8)
5. Asbestos	177	1.2 (1.0, 1.4)	41	1.6 (1.1, 2.2)	70	1.1 (0.8, 1.4)	32	1.1 (0.7, 1.5)
6. Crystalline Silica	238	1.3 (1.1, 1.5)	47	1.5 (1.1, 2.0)	109	1.4 (1.1, 1.8)	34	0.8 (0.6, 1.2)
8. Glass Dust	18	2.0 (1.1, 3.5)	2	0.9 (0.2, 3.2)	8	2.2 (1.1, 4.7)	3	1.5 (0.5, 4.3)
12. Clay Dust	28	1.9 (1.2, 3.0)	6	2.1 (1.0, 4.6)	16	2.6 (1.5, 4.4)	2	0.6 (0.2, 2.2)
15. Brass Dust	24	1.6 (1.0, 2.6)			13	1.8 (1.0, 3.3)	5	2.0 (0.9, 4.6)
21. Aluminium Alloy Dust	63	1.5 (1.1, 2.0)	12	1.5 (0.9, 2.6)	31	1.7 (1.2, 2.5)	11	1.3 (0.8, 2.4)
23. Cosmetic Talc	13	1.3 (0.7, 2.4)			6	1.3 (0.6, 2.8)		
24. Borates	11	2.0 (1.0, 3.9)	3	2.2 (0.8, 6.5)			6	5.8 (2.5, 13.4)
35. Copper Dust	47	1.3 (0.9, 1.8)	11	1.6 (0.9, 2.8)	20	1.2 (0.8, 1.9)	9	1.3 (0.7, 2.3)
36. Zinc Dust	26	1.6 (1.0, 2.6)	9	3.0 (1.5, 5.8)	12	1.8 (1.0, 3.3)	5	1.6 (0.7, 3.8)
40. Lead Chromate	35	1.1 (0.8, 1.7)	4	0.7 (0.3, 1.7)	18	1.3 (0.8, 2.2)	6	1.1 (0.5, 2.3)
53. Natural Rubber	44	1.2 (0.8, 1.7)	8	1.2 (0.6, 2.2)	19	1.1 (0.7, 1.8)	5	0.7 (0.3, 1.6)
64. Poly-Acrylates	29	1.3 (0.8, 1.9)	6	1.3 (0.6, 2.9)	14	1.4 (0.8, 2.3)	2	0.4 (0.1, 1.5)
68. Urea-Formald.	50	1.5 (1.1, 2.1)	10	1.7 (0.9, 3.1)	24	1.7 (1.1, 2.7)	8	1.1 (0.6, 2.1)
76. Soot	91	1.2 (0.9, 1.5)	14	1.0 (0.6, 1.6)	42	1.3 (0.9, 1.7)	14	0.9 (0.6, 1.5)
83. Nitrogen Oxides	240	1.6 (1.3, 1.9)	54	2.2 (1.6, 3.0)	107	1.7 (1.3, 2.1)	42	1.3 (1.0, 1.8)
85. Hydrogen Fluoride	38	2.0 (1.3, 2.9)	9	2.5 (1.3, 4.9)	17	2.1 (1.2, 3.5)	7	1.5 (0.7, 3.1)
90. Natural Gas	24	1.0 (0.6, 1.5)	7	1.7 (0.8, 3.5)	10	1.0 (0.6, 1.9)	4	0.9 (0.4, 2.1)

Table 7-15: Rate ratio estimates, from strategy 2, for selected chemicals, with respect to histological subtypes of lung cancer

	All lung cancer		Small cell		Squamous cell		Adenocarcinoma	
	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)
91. Methane	41	1.0 (0.7, 1.4)	9	1.3 (0.7, 2.4)	20	1.2 (0.8, 1.9)	7	0.9 (0.5, 1.8)
97. Coal Gas	8	0.6 (0.3, 1.2)			3	0.5 (0.2, 1.5)		
98. Gas Welding Fumes	115	1.5 (1.2, 1.8)	28	1.9 (1.3, 2.8)	48	1.4 (1.1, 1.9)	21	1.3 (0.8, 1.9)
101. Metal Oxide Fumes	190	1.3 (1.1, 1.5)	45	1.7 (1.2, 2.4)	82	1.4 (1.1, 1.7)	36	1.2 (0.8, 1.6)
104. Chromium Fumes	43	2.2 (1.5, 3.3)	12	3.2 (1.8, 5.7)	20	2.3 (1.4, 3.7)	7	1.5 (0.7, 3.1)
105. Manganese Fumes	60	1.6 (1.2, 2.2)	16	2.4 (1.4, 3.9)	24	1.5 (1.0, 2.3)	13	1.7 (1.0, 2.8)
107. Nickel Fumes	42	2.1 (1.4, 3.1)	12	3.0 (1.7, 5.4)	19	2.1 (1.3, 3.5)	7	1.5 (0.7, 3.0)
108. Copper Fumes	47	2.2 (1.5, 3.1)	9	2.1 (1.1, 4.0)	25	2.5 (1.6, 3.9)	7	1.7 (0.8, 3.4)
109. Zinc Fumes	39	1.6 (1.1, 2.3)	8	1.8 (0.9, 3.5)	16	1.4 (0.9, 2.4)	8	1.8 (0.9, 3.4)
117. Diesel Eng.Emissions	165	1.2 (1.0, 1.5)	32	1.4 (1.0, 2.0)	77	1.4 (1.1, 1.8)	24	0.9 (0.6, 1.3)
122. Propane Eng.Emiss.	28	1.7 (1.1, 2.6)	4	1.4 (0.6, 3.4)	15	1.9 (1.1, 3.3)	5	1.7 (0.7, 3.9)
127. Alkali, Caustic Solutions	72	1.3 (1.0, 1.7)	15	1.6 (1.0, 2.6)	32	1.3 (0.9, 1.9)	10	0.9 (0.5, 1.6)
156. Kerosene	69	1.6 (1.2, 2.2)	10	1.2 (0.6, 2.1)	32	1.7 (1.2, 2.5)	14	1.8 (1.0, 3.0)
158. Heating Oil	53	1.4 (1.0, 2.0)	12	1.9 (1.1, 3.3)	24	1.5 (1.0, 2.3)	12	1.7 (1.0, 3.0)
161. Cutting Fluids	85	1.3 (1.0, 1.7)	22	2.1 (1.4, 3.2)	38	1.5 (1.0, 2.1)	18	1.5 (1.0, 2.4)
163. Coal Tar and Pitch	23	1.0 (0.7, 1.7)	4	0.9 (0.4, 2.3)	10	1.0 (0.5, 1.9)	3	0.7 (0.3, 1.9)
171. Cutting Fluids post 1955	62	1.7 (1.2, 2.3)	16	2.5 (1.5, 4.1)	25	1.8 (1.2, 2.7)	14	2.0 (1.2, 3.3)
174. Inks	37	1.5 (1.0, 2.2)	7	1.4 (0.7, 2.8)	12	1.2 (0.7, 2.1)	8	1.5 (0.8, 2.9)
177. Fluorides	42	1.8 (1.3, 2.7)	9	2.1 (1.1, 4.1)	19	2.0 (1.2, 3.2)	9	1.8 (0.9, 3.3)
178. Chromium (VI) Comp.	90	1.4 (1.1, 1.8)	18	1.6 (1.0, 2.5)	41	1.5 (1.1, 2.1)	18	1.4 (0.9, 2.2)

Table 7-15: Rate ratio estimates, from strategy 2, for selected chemicals, with respect to histological subtypes of lung cancer

	All lung cancer		Small cell		Squamous cell		Adenocarcinoma	
	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)	Exp Cases	RR (90% CL)
182. Magnesium Compounds	19	1.9 (1.1, 3.4)	4	1.9 (0.7, 4.9)	9	2.2 (1.1, 4.5)	4	2.0 (0.8, 5.1)
183. Aluminium Compounds	199	1.4 (1.2, 1.7)	39	1.6 (1.2, 2.3)	95	1.6 (1.3, 2.1)	32	1.1 (0.8, 1.5)
187. Manganese Compounds	71	1.5 (1.1, 2.0)	18	2.0 (1.3, 3.3)	30	1.5 (1.0, 2.2)	16	1.6 (1.0, 2.7)
190. Nickel Compounds	79	1.7 (1.3, 2.2)	18	2.1 (1.3, 3.3)	35	1.7 (1.2, 2.4)	15	1.5 (0.9, 2.5)
191. Copper Compounds	128	1.3 (1.1, 1.6)	30	1.8 (1.2, 2.6)	60	1.5 (1.1, 2.0)	20	1.0 (0.6, 1.5)
192. Zinc Compounds	107	1.4 (1.1, 1.7)	24	1.8 (1.2, 2.7)	48	1.4 (1.0, 1.9)	19	1.3 (0.9, 2.1)
193. Arsenic Compounds	31	0.8 (0.6, 1.2)	5	0.8 (0.3, 1.7)	19	1.3 (0.8, 2.0)	6	0.8 (0.4, 1.7)
195. Cadmium Compounds	11	1.7 (0.9, 3.5)	2	1.7 (0.5, 6.2)	5	1.9 (0.8, 4.7)	3	2.4 (0.8, 7.1)
196. Tin Compounds	92	1.4 (1.1, 1.8)	18	1.4 (0.9, 2.3)	40	1.4 (1.0, 2.0)	20	1.6 (1.0, 2.4)
204. Alkanes (C5-C17)	368	1.4 (1.2, 1.7)	66	1.4 (1.0, 1.8)	169	1.6 (1.3, 2.0)	70	1.4 (1.0, 1.8)
213. Glycol Ethers	26	1.1 (0.7, 1.7)			12	1.2 (0.7, 2.1)	3	0.5 (0.2, 1.4)
214. PAH (Any)	581	1.1 (1.0, 1.4)	103	1.2 (0.8, 1.6)	261	1.4 (1.1, 1.8)	114	1.2 (0.9, 1.6)
219. Benzo(a)pyrene	220	1.2 (1.0, 1.4)	37	1.1 (0.8, 1.6)	103	1.3 (1.0, 1.7)	38	1.0 (0.7, 1.4)

7.8 Sensitivity analyses

The following sections are comprised of a number of verifications of the methods used in the regression models. The purpose of this section is to provide some means for evaluating how sensitive results were to particular assumptions and, in some cases, to give a sense of how valid certain design decisions were.

7.8.1 Semi-Bayes models

In semi-Bayes modeling, the prior variance is considered a smoothing parameter that determines how far estimates are pulled back towards the mean of the set of estimates, and care must be taken when specifying its value. I specified a tenfold range for the prior variance in the model without any exchangeability information (strategy 7) and a sevenfold range when the 31 covariates for exchangeability were added to the second level of the model (strategy 8). Analyses were carried out that assumed three other values for the prior variance: infinity (corresponding to the maximum likelihood estimator), twofold, and 0 (empirical-Bayes). Table 7-16 presents the mean of the beta logistic coefficients and the mean of the standard errors for the set of 184 parameters modeled at any level of exposure. Three chemicals were also selected as examples of different behaviours of the models under the various assumptions.

As expected, the smaller the specified value of the prior variance, the more the estimates were shrunk to the prior mean, which would be close 0.014 (or an RR of 1.01) for this set of 184 estimates, as seen in the panel of the naïve prior, under the empirical Bayes design. Examples, such as glass dust, highlight that some point estimates were indeed sensitive to the specification of the prior variance, though inference based on statistical tests would not have changed. With the naïve prior, the empirical value for the effect of exposure to glass dust shifted from 1.43 (with some shrinkage) to 1.10 (with more shrinkage); however, the associated P-value would have remained statistically non-significant. The prior variances also had an appreciably large influence on the width of the confidence intervals for all estimates. In both semi-Bayes strategies, the intervals at the twofold specification were on average half the length of the intervals at the tenfold and sevenfold specification.

Table 7-16: Sensitivity of results to various specifications of the prior variance, T^2, at ANY level of exposure, for K=184 chemicals

	MLE [a] $T^2 = \infty$ [c] Strategy 6	*Common prior (Naïve model)* [b]			*Exchangeability information* [b]		
		Semi-Bayes T^2@10fold [c] Strategy 7	Semi-Bayes T^2@2fold [c]	Empirical-Bayes $T^2=0$ [c]	Semi-Bayes T^2@7fold [c] Strategy 8	Semi-Bayes T^2@2fold [c]	Empirical-Bayes $T^2=0$ [c]
Mean beta	-0.018	0.002	0.010	0.014	0.008	0.017	0.021
Mean standard error	0.427	0.309	0.151	0.046	0.303	0.180	0.102
Selected chemicals: [d]	*RR (90% CL)*	*RR (90% CL)*	*RR (90% CL)*	*RR (90% CL)*	*RR (90% CL)*	*RR (90% CL)*	*RR (90% CL)*
Asbestos	1.01 (0.73, 1.38)	0.99 (0.74, 1.33)	0.98 (0.80, 1.20)	1.01 (0.98, 1.05)	0.97 (0.73, 1.31)	0.93 (0.72, 1.20)	0.92 (0.73, 1.16)
Glass Dust	1.61 (0.78, 3.33)	1.43 (0.81, 2.54)	1.10 (0.84, 1.45)	1.02 (0.94, 1.10)	1.49 (0.86, 2.59)	1.23 (0.91, 1.65)	1.14 (1.01, 1.29)
Silicon Carbide	0.65 (0.38, 1.10)	0.72 (0.46, 1.12)	0.92 (0.72, 1.17)	1.01 (0.95, 1.07)	0.76 (0.49, 1.18)	0.98 (0.74, 1.28)	1.13 (1.01, 1.26)

[a] MLE is maximum likelihood estimate, all chemicals in a single regression model, strategy 6.

[b] 'Common prior' refers to the strategy where all chemical effects were assumed exchangeable, whereas 'exchangeability information' refers to the strategy with information about chemical properties and previous evidence added to the second-level model.

[c] T^2, the prior variance, was set at 0.35 for a tenfold range of first-level parameter estimates, 0.25 for a sevenfold range, 0.03 for a twofold range, or 0, which indicates empirical-Bayes was used instead and the prior variance was estimated instead of specified. The hypothetical value of infinity for T^2 implies no shrinkage and thus the MLE.

[d] Three chemicals were selected to illustrate different behaviours of the estimates under the various assumptions.

A number of approaches were also considered for how to handle the covariate for previous evidence in the second-level model of strategy 8 (see step iii, 6.10). The primary approach was, for most chemicals, to set the second-level covariate to zero, signifying the null value for previous evidence, and for the few currently suspected carcinogens, to set the value to magnitudes suggested by recent meta-analyses. For sensitivity analyses, other approaches were considered: assigning values of one for all the suspected carcinogens (reducing the covariate to a dichotomous variable), or removing the covariate for previous evidence entirely. Aside from an effect on a couple of chemicals, estimates were not affected materially by these different assumptions (data not shown). For example, standard errors were unaffected down to several decimal places, and the average of the logistic beta estimates, when comparing the primary approach to the approach where previous evidence was excluded, shifted from 0.008 to 0.004.

7.8.2 Full model for strategy 6

The main results for the full model analysis (strategy 6), and consequently the semi-Bayes analyses (strategies 7 and 8), made use of a model with variables for 184 chemicals at any level of exposure. Of the 231 chemicals considered, many were eliminated for conceptual reasons (see section 6.8), but reasonable disagreement is possible as to which chemicals should have remained in this model. Table 7-17 describes the sensitivity of the maximum likelihood estimates in the full model to other regression model designs. While these models differed as to the number of chemicals included in the estimation, for comparability purposes, the descriptive statistics in the table were based on the 110 chemicals common to all four of the designs discussed below.

Compared to my preferred model, results from the model that included nearly all the chemicals simultaneously (section A, K=228 chemicals) showed only moderately similar point estimates (Spearman correlation, rho=0.81). This larger model resulted in confidence intervals that were on average 19% longer. Many of the adjustments in this model, however, would be deemed inappropriate, leading to nearly certain over-adjustment of effects (see section 6.8). Some of the chemicals whose estimates changed appreciably, included: hydrogen cyanide, whose rate ratio estimates changed from 3.2 in my preferred model to 0.9 in the larger model; creosote, with a shift from 0.62 to 0.09;

and nickel fumes, with a shift from 0.3 to 1.8. The chemicals whose estimates shifted dramatically tended to be those that occurred with low prevalence in the study population, or were highly correlated with other chemicals.

Compared to the preferred model, results from the model that deleted nearly all chemicals that could potentially pose some conceptual difficulties with the interpretation of adjusted estimates (section B, K=117 chemicals) showed similar results (Spearman correlation, rho=0.87) with few point estimates noticeably affected.

Finally, the model with deletions of one of each pair of chemicals that were highly correlated (section C, K=170 chemicals, deletions if r>0.7), possibly leading to issues of multi-collinearity in my preferred model, showed little difference in a comparison of respective point estimates (Spearman correlation, rho=0.90). Two exceptions included: nickel fumes, whose point estimate shifted from 0.3 to 1.7; and natural rubber, whose estimate shifted from 6.4 to 1.9.

Table 7-17: Sensitivity of results to various designs of the regression model for strategy 6

	Strategy 6 [a] *(K=184)*	*A. Essentially no deletion of chemicals (K=228)*	*B. Stricter deletion of chemicals even for minor issues* [b] *(K=117)*	*C. Deletion of chemicals to avoid r>0.7* [c] *(K=170)*
Mean beta [d]	0.033	0.039	0.014	0.014
Mean standard error [d]	0.431	0.514	0.376	0.404
Selected chemicals: [e]	*RR (90% CL)*	*RR (90% CL)*	*RR (90% CL)*	*RR (90% CL)*
Asbestos	1.01 (0.73, 1.38)	1.00 (0.72, 1.39)	0.94 (0.72, 1.23)	1.00 (0.73, 1.37)
Glass Dust	1.61 (0.78, 3.33)	1.41 (0.66, 3.01)	1.56 (0.81, 3.03)	1.67 (0.81, 3.41)
Silicon Carbide	0.65 (0.38, 1.10)	0.76 (0.43, 1.37)	0.66 (0.41, 1.07)	0.64 (0.38, 1.06)

[a] Strategy 6 involved logistic regression with variables for all 184 chemicals estimated simultaneously.

[b] Minor issues included possible over-adjustment of effects, such as with silica and the silica-based excavation dust.

[c] For each pair of variables with a Spearman correlation greater than 0.7, one of the pair was deleted.

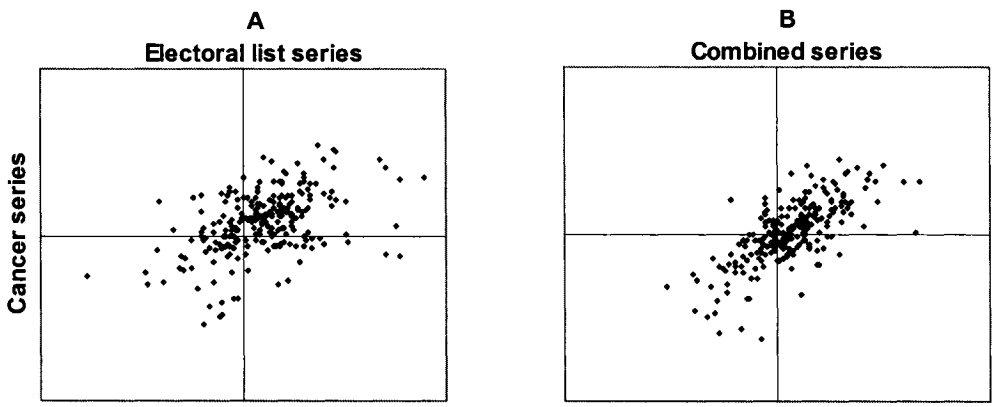
[d] For comparison purposes, the mean values were calculated on the 110 chemicals common to all four designs.

[e] Three chemicals were selected to illustrate the behaviour of the estimates under the various designs.

7.8.3 Control series options

Given the design of the study, two sources for the control series were possible: men diagnosed with cancers other than lung cancer, and men contacted through an electoral list or via random digit dialling (see section 4.2.3). Although the analyses of the thesis used the cancer patients for the control series, it was important to consider whether estimates would have been different had the other option for the controls been used. The results in Figure 7-2 compare the main estimates (cancer control series, N=2172) to the estimates using the electoral-list control series (N=533). A third option was also included (see section 6.14.2): an equal-parts combination of the cancer series and electoral-list series (N=1066). The electoral list series and cancer series options resulted in point estimates that were often quite different, near randomly distributed, as the diffuse scatter in box A demonstrates. Compared to the cancer series, the confidence limits from the electoral list design were on average 46% wider and the confidence limits from the equal-parts combined control series were 18% wider, on average.

Figure 7-2: Scatter plots comparing logistic coefficients from three different design options for the control series



The appendices provide rate ratio and confidence limit estimates for all the chemicals under the different designs of the control series.

7.8.4 Respondent status

Since some of the questionnaire respondents in the Montreal study were acting as proxies for the study subject (usually a spouse), and since proxy response may be of poorer quality for occupational history than self-response, proxy status was included as a

dichotomous variable in the statistical analyses. In order to evaluate whether proxy status indeed confounded the associations between occupational exposures and lung cancer and whether statistical adjustment was sufficient, analyses were conducted that restricted the study population to self-responders. The scatter plot in Figure 7-3 compares the estimates for all chemicals. Estimates were similar between the two approaches.

Figure 7-3: Scatter plot comparing logistic coefficients, under two options for controlling respondent status

7.8.5 Issues related to possible over-adjustment of estimates

Due to the conceptual issues with adjusting the effects of particular chemicals for the effects of others, many chemicals were not included in the larger modeling strategies. Section 6.8 discussed many of these issues and Table 6-5 lists those chemicals not included in the preferred full model design of 184 chemicals. Nevertheless, issues still remained that posed some difficulties in the interpretation of the estimated logistic coefficients. For example, compounds such as excavation dust and Portland cement were retained in the model even though they are both composed largely of silica. This may have lead to the effect of silica itself being underestimated. The results in Table 7-18 are from analyses that attempted to clarify some of these issues. In all cases, the table lists which chemicals were included in each run of the regression model, usually adding the variables for one chemical at a time, although the order of additions was arbitrary. All models included adjustment for the standard set of eight non-occupational confounders.

Since estimates appeared relatively stable across the variations of the models, conceptual issues with chemical definitions could, for the most part, be ignored when interpreting the parameter estimates in the tables of main results. However, the following exceptions should be noted:

The estimate for crystalline silica gradually shifted closer to unity with each addition to the model of a silica-composed compound. While I cannot rule out the role of confounding, this result is also consistent with the presence of over-adjustment in the larger models, in respect to the effect of silica.

Although they are distinct chemicals, natural rubber and styrene-butadiene rubber are often blended together prior to use. This functional relationship, in addition to the high correlation of the variables in the dataset (r=0.82), caused some difficulties for the interpretation of the mutually adjusted estimates. The estimates from the mutually unadjusted/adjusted approaches provided quite different pictures of the relative effects.

Chromium fumes and nickel fumes were highly correlated (r=0.97). Each appeared to have an elevated estimate when assessed individually, but when adjusted for each other, the estimates shifted radically with large increases in the estimated variance, indicating the instability of the logistic regression estimation in these simpler models, as well as in strategy 6.

Table 7-18: Sensitivity of results, for selected chemicals, with respect to issues of over-adjustment of estimates

Topic	Model [a]	Chemical (id)	Exp Cases	RR (90% CL)
Silica and various dusts				
	1	Crystalline Silica (6)	238	1.3 (1.1, 1.5)
	2	Crystalline Silica (6)	238	1.2 (1.0, 1.5)
		Abrasives Dust (1)	237	1.1 (1.0, 1.4)
	3	Crystalline Silica (6)	238	1.1 (0.9, 1.3)
		Abrasives Dust (1)	237	1.2 (1.0, 1.4)
		Excavation Dust (3)	109	1.4 (1.0, 1.8)
	4	Crystalline Silica (6)	238	1.0 (0.8, 1.3)
		Abrasives Dust (1)	237	1.2 (1.0, 1.4)
		Excavation Dust (3)	109	1.4 (1.0, 1.8)
		Portland Cement (7)	79	1.2 (0.9, 1.6)
	5	Crystalline Silica (6)	238	1.0 (0.8, 1.3)
		Abrasives Dust (1)	237	1.2 (1.0, 1.4)
		Excavation Dust (3)	109	1.4 (1.0, 1.8)
		Portland Cement (7)	79	1.3 (0.9, 1.7)
		Brick Dust (11)	34	0.8 (0.5, 1.2)
	6	Crystalline Silica (6)	238	1.0 (0.8, 1.2)
		Abrasives Dust (1)	237	1.2 (1.0, 1.4)
		Excavation Dust (3)	109	1.4 (1.1, 1.9)
		Portland Cement (7)	79	1.3 (1.0, 1.8)
		Brick Dust (11)	34	0.8 (0.5, 1.2)
		Clay Dust (12)	28	1.9 (1.2, 3.1)
	7	Crystalline Silica (6)	238	1.0 (0.8, 1.3)
		Abrasives Dust (1)	237	1.2 (1.0, 1.4)
		Excavation Dust (3)	109	1.5 (1.1, 2.0)
		Portland Cement (7)	79	1.3 (1.0, 1.8)
		Brick Dust (11)	34	0.8 (0.5, 1.2)
		Clay Dust (12)	28	1.9 (1.2, 3.0)
		Concrete Dust (13)	97	0.9 (0.7, 1.3)
Glass dust and fibres				
	1	Glass Dust (8)	18	2.0 (1.1, 3.5)
	2	Glass Fibres (9)	50	0.9 (0.7, 1.2)

[a] Model numbers refer to separate regression models, each including only the chemicals listed to the right, with adjustment for the standard eight non-occupational confounders.

Table 7-18: Sensitivity of results, for selected chemicals, with respect to issues of over-adjustment of estimates

Topic	Model [a]	Chemical (id)	Exp Cases	RR (90% CL)
	3	Glass Dust (8)	18	2.0 (1.1, 3.5)
		Glass Fibres (9)	50	0.9 (0.7, 1.2)
Rubber				
	1	Natural Rubber (53)	44	1.2 (0.8, 1.7)
	2	Styrene-Buta.Rubber (70)	38	0.9 (0.6, 1.3)
	3	Natural Rubber (53)	44	2.4 (1.3, 4.4)
		Styrene-Buta.Rubber (70)	38	0.4 (0.2, 0.8)
Engine emissions				
	1	Gas Eng.Emissions (115)	379	0.9 (0.8, 1.1)
	2	Diesel Eng.Emissions (117)	165	1.2 (1.0, 1.5)
	3	Jet Fuel Eng.Emiss. (121)	3	0.4 (0.2, 1.3)
	4	Propane Eng.Emiss. (122)	28	1.7 (1.1, 2.6)
	5	Gas Eng.Emissions (115)	379	0.9 (0.8, 1.0)
		Diesel Eng.Emissions (117)	165	1.3 (1.0, 1.5)
	6	Gas Eng.Emissions (115)	379	0.9 (0.8, 1.0)
		Diesel Eng.Emissions (117)	165	1.2 (1.0, 1.5)
		Jet Fuel Eng.Emiss. (121)	3	0.5 (0.2, 1.4)
	7	Gas Eng.Emissions (115)	379	0.9 (0.8, 1.0)
		Diesel Eng.Emissions (117)	165	1.2 (1.0, 1.5)
		Jet Fuel Eng.Emiss. (121)	3	0.4 (0.2, 1.3)
		Propane Eng.Emiss. (122)	28	1.7 (1.1, 2.7)
	8	Carbon Monoxide (80)	478	0.9 (0.7, 1.2)
		Nitrogen Oxides (83)	240	2.3 (1.7, 3.0)
		Sulphur Dioxide (86)	144	0.5 (0.4, 0.7)
		PAH (Any) (214)	581	1.1 (0.9, 1.3)
		Benzo(a)pyrene (219)	220	1.1 (0.9, 1.4)
		Soot (76)	91	1.0 (0.8, 1.4)

[a] Model numbers refer to separate regression models, each including only the chemicals listed to the right, with adjustment for the standard eight non-occupational confounders.

Table 7-18: Sensitivity of results, for selected chemicals, with respect to issues of over-adjustment of estimates

Topic	Model [a]	Chemical (id)	Exp Cases	RR (90% CL)
	9	Gas Eng.Emissions (115)	379	0.8 (0.6, 1.0)
		Diesel Eng.Emissions (117)	165	1.1 (0.9, 1.4)
		Jet Fuel Eng.Emiss. (121)	3	0.4 (0.1, 1.2)
		Propane Eng.Emiss. (122)	28	1.7 (1.1, 2.7)
		Carbon Monoxide (80)	478	1.1 (0.8, 1.4)
		Nitrogen Oxides (83)	240	2.1 (1.6, 2.8)
		Sulphur Dioxide (86)	144	0.5 (0.4, 0.7)
		PAH (Any) (214)	581	1.1 (0.9, 1.4)
		Benzo(a)pyrene (219)	220	1.1 (0.8, 1.4)
		Soot (76)	91	1.1 (0.8, 1.5)
Polycyclic aromatic hydrocarbons				
	1	PAH (Any) (214)	581	1.1 (1.0, 1.4)
	2	Benzo(a)pyrene (219)	220	1.2 (1.0, 1.4)
	3	PAH (Any) (214)	581	1.1 (0.9, 1.3)
		Benzo(a)pyrene (219)	220	1.2 (1.0, 1.4)
	4	PAH (Wood) (216)	40	0.9 (0.6, 1.3)
		PAH (Petroleum) (217)	561	1.1 (0.9, 1.3)
		PAH (Coal) (218)	84	1.3 (1.0, 1.7)
		PAH (Other) (215)	187	1.2 (1.0, 1.4)
	5	PAH (Any) (214)	581	0.9 (0.6, 1.5)
		Benzo(a)pyrene (219)	220	1.1 (0.9, 1.4)
		PAH (Wood) (216)	40	0.9 (0.6, 1.3)
		PAH (Petroleum) (217)	561	1.1 (0.7, 1.8)
		PAH (Coal) (218)	84	1.3 (0.9, 1.7)
		PAH (Other) (215)	187	1.2 (1.0, 1.4)
Natural gases				
	1	Natural Gas (90)	24	1.0 (0.6, 1.5)
	2	Coal Gas (97)	8	0.6 (0.3, 1.2)
	3	Hydrogen (79)	19	0.8 (0.5, 1.5)
		Carbon Monoxide (80)	478	1.2 (1.0, 1.4)
		Hydrogen Sulphide (87)	37	1.0 (0.7, 1.5)
		Methane (91)	41	1.0 (0.7, 1.5)

[a] Model numbers refer to separate regression models, each including only the chemicals listed to the right, with adjustment for the standard eight non-occupational confounders.

Table 7-18: Sensitivity of results, for selected chemicals, with respect to issues of over-adjustment of estimates

Topic	Model [a]	Chemical (id)	Exp Cases	RR (90% CL)
	4	Natural Gas (90)	24	0.9 (0.4, 1.7)
		Coal Gas (97)	8	0.4 (0.1, 1.1)
		Hydrogen (79)	19	1.3 (0.6, 2.5)
		Carbon Monoxide (80)	478	1.2 (1.0, 1.4)
		Hydrogen Sulphide (87)	37	1.0 (0.7, 1.4)
		Methane (91)	41	1.3 (0.7, 2.3)
High correlation between chromium and nickel fumes				
	1	Chromium Fumes (104)	43	2.2 (1.5, 3.3)
	2	Nickel Fumes (107)	42	2.1 (1.4, 3.1)
	3	Chromium Fumes (104)	43	4.1 (0.6, 27.2)
		Nickel Fumes (107)	42	0.5 (0.1, 3.7)
Dusts and fumes				
	1	Copper Dust (35)	47	1.3 (0.9, 1.8)
	2	Copper Fumes (108)	47	2.2 (1.5, 3.1)
	3	Copper Dust (35)	47	1.0 (0.7, 1.5)
		Copper Fumes (108)	47	2.1 (1.4, 3.1)
	4	Iron Dust (33)	37	1.1 (0.7, 1.6)
	5	Iron Fumes (106)	94	1.4 (1.1, 1.8)
	6	Iron Dust (33)	37	1.0 (0.7, 1.4)
		Iron Fumes (106)	94	1.4 (1.1, 1.8)
	7	Zinc Dust (36)	26	1.6 (1.0, 2.6)
	8	Zinc Fumes (109)	39	1.6 (1.1, 2.3)
	9	Zinc Dust (36)	26	1.4 (0.8, 2.2)
		Zinc Fumes (109)	39	1.5 (1.0, 2.2)

[a] Model numbers refer to separate regression models, each including only the chemicals listed to the right, with adjustment for the standard eight non-occupational confounders.

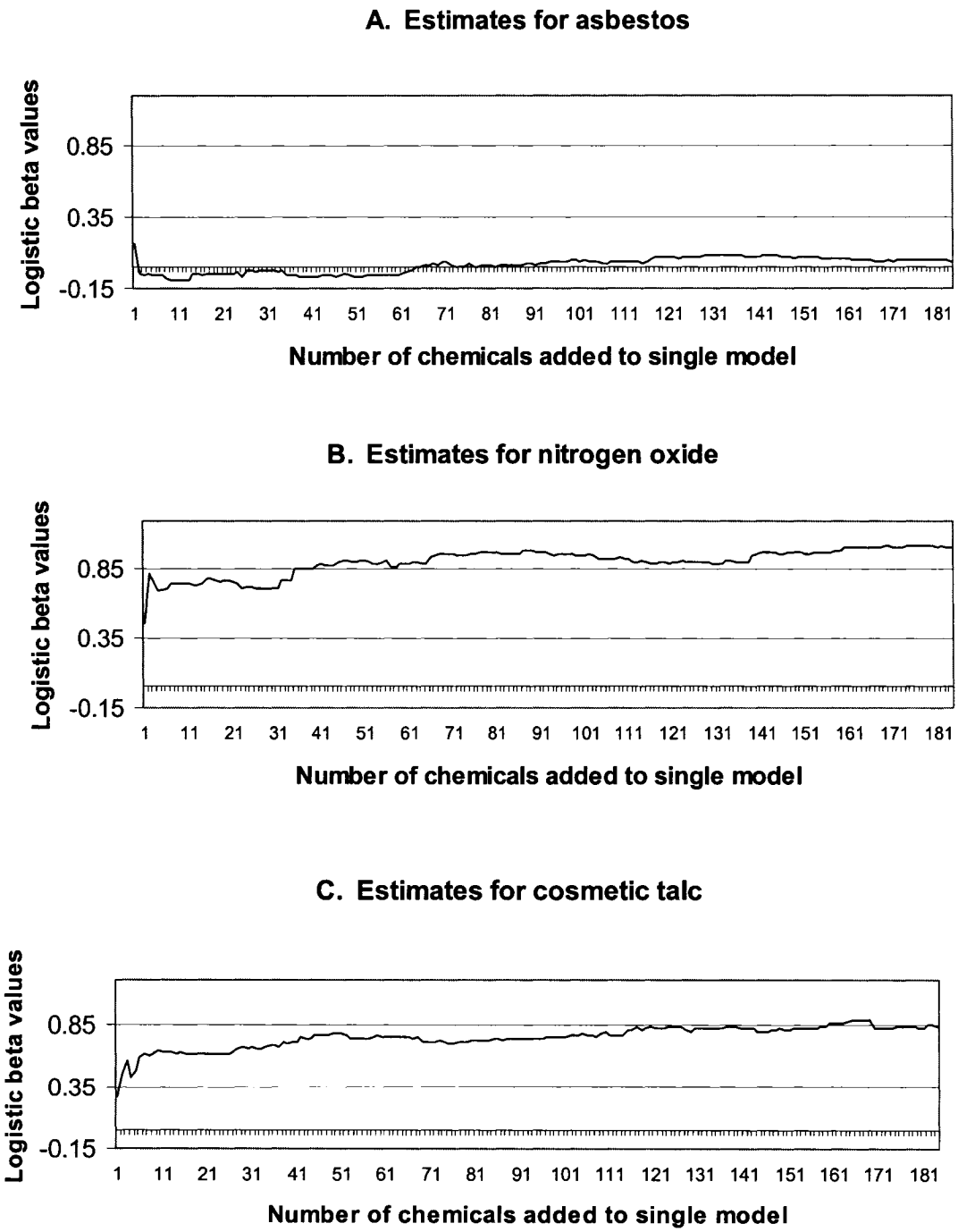
7.8.6 Issues related to correlation and confounding

This section explores how the estimates for particular chemicals change with the gradual addition of other chemicals, up to the size of the full model strategy of 184 chemicals. A few chemicals were chosen for illustrative purposes.

The estimates for asbestos are displayed in box A of Figure 7-4. The x-axis lists the number of chemicals included in the model at that part of the plot, where the leftmost value represents the model for the unadjusted effect of asbestos. Fluctuations in the estimate for a chemical can reflect any combination of the following: random deviations, adjustment for confounding, combined impact of minor amounts of confounding, minor amounts of over-adjustment, and small sample bias.

Even this small sample of three chemicals illustrated that there was no single pattern which described all chemicals. While the important changes to each of the chemical's estimates occurred in the left extreme of the graph, that is, after adjusting for a few important confounders, it could not be predicted with certainty that the inclusion of more covariates would have had the effect of decreasing or increasing the RR estimates. These phenomena seem to be idiosyncratic and depend on the complex inter-correlations among the chemical exposures.

Figure 7-4: Tracking the logistic beta estimate for a chemical across models with the gradual addition of variables for all other chemicals



8 Discussion

8.1 Limitations of the study and of the analyses

When interpreting the results of the present thesis, several considerations must be borne in mind: suspected and expected study biases, issues with exposure assessment, and limitations in the design of the analyses.

8.1.1 Selection of study subjects

The case series of lung cancer patients would have represented a near-complete ascertainment of cases arising in metropolitan Montreal, but patients with lung cancer were excluded from the study for three non-contiguous years of the six years of recruitment (Siemiatycki, 1991). The case series can nevertheless be considered a random sample of lung cancer cases over the years of the study, as any systematic difference in characteristics of cases arising in different years is very unlikely.

The control series of cancer patients, representing near-complete ascertainment of the respective types of cancer from the Montreal-area hospitals, were restricted to residents of metropolitan Montreal. The electoral-list control series was sampled from an enumeration of the Quebec population, and also restricted to residents of metropolitan Montreal (Siemiatycki, 1991). Thus, both options for the control series were consistent with sound epidemiologic design (Rothman and Greenland, 1998), and in theory could be considered equivalent for representing the distributions of exposures in the source population (Miettinen, 1999). However, the scatter plot in Figure 7-2, comparing estimates using the cancer series to estimates using the electoral-list series, suggests that in practice there were many discrepancies between the pairs of estimates.

The response rates of the cancer series were, on average, higher than the response rate of the electoral-list series. The average response rate of the cancer series was 81.5%, ranging from 63% to 92% for the different subtypes of cancer, while 72% of the men responded who were contacted on the electoral list or from random digit dialling. Insofar as low response rates can induce unrepresentative or biased samples, the cancer series would have been less prone to such a bias.

8.1.2 Bias from choice of cancers in the control series

It would have been desirable to remove from the control series, for each chemical of interest, those cancers known to be associated with the chemical (Rothman and Greenland, 1998). For example, in the assessment of aromatic amines and lung cancer, removing bladder cancer would have been justified (Ward et al., 1991; Markowitz and Levin, 2004). However, because of the lack of any relevant knowledge about the potential carcinogenic effects of the majority of the chemicals in the database, it was utterly unfeasible to conduct such a tailored set of analyses. Instead, an attempt was made to design a cancer control series which would be acceptable for nearly all the chemicals in the study. This was done by restricting the number of subjects with any of the subtypes of cancer, so that no single cancer site accounted for more than 10% of the total cancer controls. However, bias was still a possibility if certain chemicals were strongly related to any of the cancer types in the control series. An example of a plausible magnitude of this bias was given in section 6.3.

8.1.3 Confounding bias

Twenty variables were used to represent the eight non-occupational confounders that were adjusted for throughout the analyses. The extent to which these eight 'confounders' actually brought about confounding of the chemical estimates was not assessed, as it would be different for every chemical.

Of the potential non-occupational confounders considered for adjustment in the models, two suspected determinants of the risk of lung cancer were not included: family history (Shaw et al., 1991) and diet (Byers, 1994). Family history of lung cancer was not collected in the Montreal study. However, it is unlikely that it would be even moderately correlated with exposure to particular chemicals. The non-occupational characteristics, in general, would more likely be correlated with employment in occupations, and only indirectly with particular chemical exposures. Questions on diet were included in the interviews and questionnaires of the Montreal study, but the questions were not comprehensive and it would be expected that the data on particular diet items would be of questionable quality. Some of this data was previously translated into an index of beta-carotene consumption, but too many study subjects were missing data for this variable to

be included in the models. As with family history, any confounding of the chemical effects due to diet would likely be minor.

Despite my attempts to adjust the effects of all the chemicals for mutual confounding in strategies 3 through 8, a number of considerations would suggest that residual confounding among the chemicals was still present. The foremost reason is the crude nature of representing exposures in the models. A single variable, corresponding to an ever/never dichotomization of lifetime work history, might not offer thorough control for the entire effect of an important confounding risk factor. Added to this, measurement error in the exposure variables would also leave some residual confounding in the models that purported to adjust for the effects of other chemicals (Armstrong, 1998).

8.1.4 Quality of data

Diagnoses of cancers were obtained through the pathology departments of all the participating hospitals of the Montreal study (Siemiatycki, 1991). Each case was eligible for the study only following histological confirmation. The diagnosis of lung cancer is not considered to be very difficult or error-prone, and so it is unlikely that there would be many false-positives among the lung cancer cases included in the study. There may have been under-ascertainment of true lung cancers that were not histologically confirmed. Misclassification among the histological subtypes of lung cancer was more likely, because at the time of study recruitment the histological designation was more difficult and occasionally ambiguous (Bruce Case, personal communication). Such errors may be expected to be nondifferential with respect to any of the chemical exposures assessed, so that any resulting bias of the estimates would tend to be toward the null.

Another concern relates to the validity of occupational exposure information obtained from proxy respondents, which accounted for 20% of the completed interviews among the cancer patients. Spouses and other relatives would provide less information about the work history of the study subject, and this would have the effect on the chemists' expert exposure assessment of shifting it closer to that of a traditional job-exposure matrix. This in turn could introduce exposure misclassification. However, in a sensitivity analysis, adjusting for respondent status provided similar results to restricting the study to self-responders (see section 7.8.4).

The occupational hygienists coded exposures with a consensus process that should have minimized exposure misclassifications (Siemiatycki, 1996). Re-codes have occurred over the years for particular groups of substances, and this also presumably improved the accuracy of the exposure codes. Based on the chemists' coding of exposure certainty, the regression models used in this thesis also included two variables for each chemical to distinguish uncertain and certain exposures. Further, information about the disease status of study subjects was not made available to the chemists. Given these precautions, to the extent that misclassification might have still occurred, it would have resulted in bias toward the null due to non-differential misclassification.

In contrast, if non-differential misclassifications of different exposures are correlated, then this can unexpectedly lead to bias away from the null (Chavance et al., 1992; Kristensen, 1992). While correlated errors could be expected for determinants like smoking history and alcohol use, it would be minimized in the exposure assessment of the occupational chemicals. This is because the interview information was filtered several times through several occupational hygienists, who also incorporated information external to the study in their assessments, such as from consultations with experts and from bibliographic sources (Siemiatycki, 1991).

8.1.5 Issues related to the analytic strategies

Definition of exposure variables

The analyses of chemicals were carried out at two levels of exposure, any and substantial, with the latter being a subset of the former. The designation of 'exposed at any level' was based on the comparison to exposure in the general environmental, but it had no clear definition in terms of units of concentration per volume. Likewise, the designation of 'substantial exposure' was not based on absolute exposure levels, but was rather a relative construct. In general, benchmarking would be possible by considering the occupations common to these levels in the Montreal population. This, however, would be an imperfect algorithm and would not provide a solution in the present thesis for interpreting point estimates, which relate to cumulative exposures across separate jobs. The point estimates reported in the thesis, then, must be interpreted only on a relative scale and only

qualitatively, with 'any' exposure corresponding to an average of lower concentrations and shorter durations than 'substantial' exposure.

The algorithm used to combine the different characteristics of exposure into a single index relied on somewhat arbitrary assumptions about the relative importance of concentration versus duration. More finely crafted levels of exposure, beyond the present use of 'any' and 'substantial', would require even stronger assumptions about how to weight low concentration/long duration exposures versus high concentration/short duration exposures. I sidestepped this issue by avoiding defining low and moderate exposure levels, especially as these definitions would have had to suffice for all 231 chemicals investigated in the thesis. Besides, many of the chemicals occurred with too low a prevalence to have supported finer categories in the estimation.

Without better knowledge of the induction and latency times (Rothman, 1981) for each chemical-lung cancer association, a simple approach was used by defining a single five-year period prior to diagnosis that was deemed irrelevant to the onset of the disease. Accordingly, the five years of exposure before diagnosis or interview were discounted. The standard use of five years would have introduced error due to misclassification if the minimum induction period was much longer (Rothman, 1981), as is certainly the case for some occupational substances. In the secondary analyses, the designation of the mutually exclusive time windows of 5 to 20 years and 20+ years before diagnosis was entirely arbitrary. Far more sophisticated approaches are available to analyze lag periods (Rachet et al., 2003). The analysis of exposures in the time window of 20 or more years prior to diagnosis often resulted in imprecise estimates. The results were nevertheless reported because rule-out evidence of carcinogenicity cannot be provided by only analyzing exposure periods close to diagnosis, especially when there is uncertainty about the true latency/induction period (IARC, 2004a).

Exchangeability in the semi-Bayes model

The accuracy of the results from the semi-Bayes models is dependent on the assumptions made about the exchangeability of groups of chemical effects (Greenland, 2000a). The categories of exchangeability developed for this project were based on the chemical and physical properties of the occupational chemicals. If the 30 categories specified in the

semi-Bayes model did not reflect the true nature of the carcinogenicity of the occupational chemicals, then the accuracy of the results of the semi-Bayes models would be affected. For example, a category called 'solvents' was discarded early in the process of designing the categories of exchangeability. This was because many different types of chemicals are considered solvents, and there would not be common properties suggesting that they would have similar effects on lung cancer. Including the category of solvents in the semi-Bayes models might have shrunk the estimates of these different chemicals to an inappropriate common mean. However, almost all of the chemicals assessed in the thesis were expected to have relatively small magnitudes of effect on lung cancer, which in turn implies that any misspecification of exchangeability would not have resulted in large degrees of error.

There were other characteristics of the occupational chemicals that might ideally have been included in the model. For example, particle size would be predictive of the respirability of dusts, and it would plausibly be related to their carcinogenicity to lung tissue. Many of the chemicals, however, occurred in a variety of industrial processes at different temperatures and with a different spectrum of adsorbed chemicals, thus making it difficult to attribute a single value for particle size to a given chemical.

8.2 Methodological findings

8.2.1 The modeling strategies

Although several modeling strategies were implemented in the thesis, most were used for comparative purposes only. Strategies 1 and 2, for example, were too simplistic for a study of multiple occupational exposures, but they were nevertheless included to observe whether they resulted in appreciably different results when compared to the more complex analyses. Because of the very large number of results generated by all these models, chemical-specific estimates from several of the strategies were not shown. Nevertheless, general comments can be made about some of the benefits and expected problems of each approach.

Strategy 1 involved adjusting the effect of each chemical for age only. This resulted in over 20% of the chemicals identified as 'statistically significant', but this finding could

largely be attributed to ignoring confounding from non-occupational and occupational risk factors.

Strategy 2 involved adjusting the effect of each chemical only for eight non-occupational confounders, which involves the explicit assumption that no confounding occurred among the chemicals. Given the presence of some high correlations among the exposures, several of these estimates could represent over-estimates.

Strategy 3 assessed each chemical in a separate regression model, with adjustment for the standard eight non-occupational confounders and for seven suspected lung carcinogens. Considering the ensemble of chemicals assessed in the thesis, the mean of the logistic beta estimates shifted closer to zero compared to the estimates of strategies 1 and 2. This would arguably represent a more reasonable distribution of estimates if it is expected that most of the chemicals among the 231 assessed in the thesis would *not* be lung carcinogens. Furthermore, little was lost in terms of precision by adding the seven suspected lung carcinogens over what was included in the models of strategy 2, as the standard errors were similar. Strategy 3 also had the benefit of avoiding the automaticity and arbitrary criteria of strategies 4 and 5 in the selection of confounders. Choosing which chemicals to include for confounding purposes drew on expert opinion and previous research findings, which is desirable since knowledge of subject matter should be an important input to regression modeling (Robins and Greenland, 1986). This approach also involved the recognition that either too few or too many covariates can harm estimation (Starr et al., 1986).

On the other hand, a philosophical objection to all the approaches involving K models for K exposures is that they can result in mutually contradictory models. For example, a chemical will be included in the model when it is the focus of estimation, and the estimated logistic beta can be any value from negative infinity to positive infinity. But in the next model, when it is being treated as a potential confounder, its estimated effect would be set to exactly zero if it were deleted from the model. This objection takes on relevance when the objective of a study is to estimate the separate effects of each of a set of exposures. Whether this would have practical implications in the estimation would depend on the complex pattern of correlations among the variables.

Strategy 4 involved a separate model for each chemical being assessed, and other chemicals were added to this model as confounders using an automatic forward selection with a P-value ≤ 0.25 criterion for entry. While approximately 50 chemicals were included as confounders in each of the separate regression models, on average the estimated variances were similar to or even higher than the variances from the semi-Bayes models, which included 184 chemicals in a single model. Strategy 4 offered little over the other approaches. The criterion for including a chemical as a confounder was based on a P-value, which has been widely criticized as an inappropriate indicator of the extent of confounding (Greenland, 1989b; Weinberg, 1993; Hernberg, 1996; Nurminen, 1997; Hernan et al., 2002). From Table 7-6, the failure of the P-value-based criterion to identify some important confounders might be one explanation for why so many more chemicals had elevated estimates in strategy 4 compared to strategy 3, especially in the analyses at any level of exposure. An alternative approach to the P-value criterion would be to limit covariates to those that bring about a predetermined change in the regression coefficient of the main exposure, such as a 10% relative change (Mickey and Greenland, 1989; Maldonado and Greenland, 1993). This latter strategy was not incorporated in the present thesis because it is ill-suited to assessing studies of hundreds of exposures.

Strategy 5 was the first of the strategies to attempt the estimation of the chemicals in a single model. While all 184 chemicals were eligible for the model, the only chemicals included were those that had a statistically significant point estimate with P-value ≤ 0.25. The untenable idea behind this approach was that if a chemical's estimate was not statistically significant, then it in fact was truly equal to unity (which was operationalized by excluding the chemical from the model). Such an approach to assessing a set of exposures introduces discontinuities into the estimation by setting several effects to the null value without confidence limits. If the objective of a study is to estimate the effects of K chemicals, then such an approach to modeling represents far too much certainty about the majority of the effects. In comparison, the semi-Bayes approach only shrinks estimates *toward* the prior mean, which in this study would be close to the null value, and it provides a confidence interval for each estimate.

Strategy 6 involved a single logistic regression model with variables for all 184 chemicals. Using a so-called full model has been advocated by theoreticians, though

there is recognition that such models would rarely have sufficient data to be properly fit (Miettinen, 1985). Suggestions for modeling often begin with the full model and follow with a backwards deletion of unnecessary variables using a P-value criterion for exit (Sun et al., 1996). However, one of the reasons for using the semi-Bayes models was to estimate the entire panel of exposures and avoid dropping variables from the model. In my own analyses, the RR estimates from strategy 6 were often large and imprecise, which was expected because of the small ratio of lung cancer cases to exposure variables in the logistic regression. The full model was used in the present thesis solely as an interim step to the semi-Bayes approaches.

Strategy 7 involved the same single model as in strategy 6, but now a semi-Bayes prior was specified that assumed that the effects of all the chemicals were exchangeable. There were far fewer statistically significant estimates in the semi-Bayes model results than in any of the non-Bayesian strategies. Curiously, no estimates were statistically elevated with substantial exposures, although this could partly be explained by the imprecise nature of all the analyses at that level, thus involving extensive shrinkage of estimates. In this thesis, the assumption of complete exchangeability of all effects was neither plausible nor advisable (Greenland, 2000a), and it might have led to inappropriate shrinkage of point estimates. For example, the estimates of the effects of chromium VI compounds tended to be lower in strategy 7 than when they were separated into a category of exchangeability in strategy 8.

Strategy 8 was also based on a single logistic regression model of 184 chemicals, but now with a semi-Bayes prior that grouped the chemicals by the similarity of their chemical properties. It also involved incorporating subject-matter knowledge in terms of which chemicals could be included simultaneously in a regression (like strategy 6). Introducing the information about the chemical properties influenced very few of the estimates compared to the simpler approach of strategy 7. Results were also insensitive with respect to how the second-level covariate, that represented previous evidence, was represented in the models. Even entirely excluding the covariate for previous evidence did not materially affect the results (see section 7.8.1). These results highlight the uncertainty of current knowledge about what determines an occupational chemical's carcinogenicity.

Summary of noteworthy influences of the modeling strategies

The present results included several comparisons of the distributions of estimates derived from the different models. There were several noteworthy patterns that emerged from these comparisons. In terms of the greatest influences in the various strategies, comparing the small-model approach (strategy 3) to the large-model approach (strategy 6) resulted in often dissimilar estimates. Incorporating the simplistic semi-Bayes prior (strategy 7) over the large-model, maximum likelihood approach (strategy 6), also greatly influenced the estimates in terms of shrinking large and imprecise estimates to lower, more reasonable values. On the other hand, the introduction of expert opinion in terms of the categories of exchangeability based on chemical and physical properties (strategy 8), had only a negligible influence on nearly all estimates when compared to naïve shrinkage to a common prior mean (strategy 7). In fact, de Roos et al (2001) also found, in their occupational cancer study, that including second-level covariates for crudely defined chemical properties only somewhat changed estimates compared to models with a simple prior. They found that where the magnitude of estimates changed, typically if the substance belonged to several categories of chemical properties, a loss of precision also occurred. Presumably, these estimates would nevertheless be more accurate than those of maximum likelihood (Greenland, 1993).

The particular circumstance of chromium fumes and nickel fumes would be a good example in which to explore the properties associated with some of these modeling strategies. Almost everyone exposed to nickel fumes was also exposed to chromium fumes. Therefore, it would be very difficult to separate their effects with the present data. Estimates for both of these chemicals were statistically significantly greater than 1.0 in the results from strategy 3 (chromium fumes, 2.0 (1.2, 3.2), and nickel fumes, 1.7 (1.1, 2.8)), which examined them separately and without mutual adjustment. However, because of the near-collinearity between them (r=0.97), the two estimates may be almost completely confounded. By including the two exposures in a single model (strategy 6), the collinearity drove the respective estimates to unreasonably large and imprecise values: 7.5 (0.8, 69.2) for chromium fumes, and 0.2 (0.0, 2.4) for nickel fumes. This is a pattern observed with near-collinear variables (Kleinbaum et al., 1998). In contrast, strategy 7 resulted in both estimates being shrunk to somewhat more similar values, both of which

were closer to unity (respectively, 1.4 (0.7, 2.9) and 1.1 (0.5, 2.3)). This dramatic shrinkage occurred because of the very high standard errors. In this sense, strategy 7 accounted for near-collinearity while still allowing for the mutual adjustment of the effects of these two chemicals. Finally, the assumptions of exchangeability in strategy 8 included separate categories for nickel compounds and chromium compounds. Based on these assumptions, which were used to bolster the estimates of all the chemicals when data was sparse, the estimate for chromium fumes was elevated further to 1.9 (0.6, 6.0), but the estimate for nickel fumes remained close to unity, 0.8 (0.3, 2.5). This occurred partly because other chromium compounds also had elevated estimates, influencing the estimate of the effect of chromium fumes.

Preference in certain modeling strategies

Of the eight main approaches to modeling, strategy 3 and strategy 8 represented what appear to be, respectively, the most common and the most preferred (Greenland, 1993) strategies for analyzing a set of multiple exposures. For these reasons, the results of both strategies were used to select those chemicals that would be followed-up with further analyses. Comparison of the distributions of estimates resulting from these analytic approaches showed that the corresponding point estimates were often quite different. Accordingly, they led to prioritizing entirely different chemicals in the section on ranking and selection (see section 7.6).

In the situation where one wishes to prioritize chemicals or to make inferences on a set of multiple exposures, empirical Bayes models have been widely advocated over conventional strategies (Thomas, 1985; Greenland and Robins, 1991), especially if there is a high cost to following false leads (Steenland et al., 2000). Of course, the re-ordering of parameters using empirical Bayes methods is driven by the exchangeability assumptions, and thus relies on their validity (Morris, 1983b). In the present study, the specification of a single large model was complicated by the overlapping definitions of many chemicals. This included 'chemicals' that actually represented groups of chemicals (like arsenic compounds) and chemicals that were complex mixtures of other chemicals (like gasoline exhaust). In the present thesis, the assessment of the semi-Bayes models attempted to distinguish two separate issues: the nature of the mutual adjustment of the

effects of many chemicals in a single model, and the nature of the semi-Bayes shrinkage of multiple estimates. Further comments on each of these two issues are found below.

8.2.2 Comments on mutually adjusting the effects of occupational chemicals

In designing the full model with the independent effects of 184 chemicals estimated simultaneously, an even larger model and, alternatively, a smaller model were considered as the best approach for strategy 6 (see section 6.14.2). Several questions were pertinent in arriving at the number of 184: do any of the reported estimates have complicated interpretations due to other mutual adjustments? do any of the mutual adjustments lead to potential over-adjustment? is it logically inappropriate to adjust the effects of this chemical for the effects of that one? is there interest in distinguishing the effects of any particular sets of chemicals? The answers to these questions, along with a certain amount of reasoning and logic, were used to arrive at the final model proposed for strategy 6. For example, should the effects of gasoline exhaust be adjusted for the effects of diesel exhaust, even though they have a similar profile of emissions? And furthermore, should these effects be adjusted for the effects of the emissions themselves, like carbon monoxide and nitrogen oxide? This process of reasoning was not based on statistical criteria, but rather on subject-matter knowledge. The rate ratio estimates in the tables of results (such as in Table 7-11) needed to be easily interpretable, but that requirement was balanced against the need to estimate unconfounded effects and effects of the relative contributions of these chemicals, especially in the face of multiple high correlations.

Inasmuch as the full model (strategy 6) involved a mutual adjustment of the effects of individual chemicals, it must be borne in mind that this model is a theoretical construct that involves many assumptions. For example, it is not logically possible to consider the effect of, say, whole diesel exhaust, while holding fixed the exposure levels of many of its constituent emissions, and yet, in principle, this is what the fully-adjusted model was intended to accomplish. Thus, the point estimates for whole diesel exhaust correspond to the theoretical *remaining* effect of diesel exhaust after having removed the effects of individual emissions. While biases of over-adjustment were unavoidable to some extent, the ones that were identified appeared to have only rarely made a difference in terms of the numerical estimates, as shown in the sensitivity analyses (see section 7.8.5). The

over-adjustment of estimates poses conceptual difficulties nonetheless when interpreting the results of a few identified sets of chemicals, such as silica-based exposures and engine exhausts.

8.2.3 Further comments on the semi-Bayes models

Semi-Bayes and empirical Bayes models have been strongly advocated for analyzing multiple exposures (Thomas et al., 1985; Greenland, 1992) Part of the argument supporting such models is based on the concern that in a model like that of strategy 3, the opinions regarding whether a particular chemical should be included as a confounder or not, are represented with too much certainty. The model in strategy 6, and the models based on it (strategies 7 and 8), avoided pre-selecting which chemicals confounded each other by estimating all their effects simultaneously, and it did this within a Bayesian framework. On the other hand, the size of the model, the risk of over-adjusting several effects, the difficulties of specifying a meaningful Bayesian prior, and unfamiliarity with the software necessary for Bayesian modeling would naturally lead to hesitation about using this modeling approach. Empirical Bayes models, in general, sit uneasily between Frequentist methods and Bayesian methods. Largely due to criticisms of the arbitrariness of the different estimation techniques used in empirical-Bayes analyses (Robert, 2001), pure Bayesians have also criticised these approaches. For example, Lindley (1983) has written that "no one is less Bayesian than an empirical Bayesian." But others have sought to clarify the commonalities between Frequentist and Bayesian models under the framework of multilevel modeling (Greenland, 2000a).

Large numbers of independent variables

Shortly after the introduction of the maximum likelihood logistic function, cautions in its use and misuse began to appear (Gordon, 1974). Multiple logistic regression was developed as a tool for handling several independent variables in relatively sparse data (Walker and Duncan, 1967), especially in circumstances that stratified analyses, like the Mantel-Haenszel, could not handle. Gordon (1974) suggested that in any application of multiple logistic regression, the results should be explored by reverting to stratified analyses. However, in the context of the present thesis, it would be impossible to verify the regression estimates in this fashion. For example, the logistic regression (strategy 6)

analysis of the 184 chemicals at any level of exposure could be thought of in terms necessary for a Mantel-Haenszel analysis. This would require the following numbers of dichotomous variables: one for the dependent variable, 184 for the exposures of interest, 125 to account for uncertain levels of exposure, and 20 for the non-occupational confounders. This would be equivalent to $2^1 2^{184} 2^{125} 2^{20} \approx 2.18 \times 10^{99}$ cells of data. Gordon (1974) goes on to write that even with a modest amount of sample data, it is practically impossible to test the appropriateness of the assumptions underlying logistic regression, and we must "rely on other evidence (or our hopes) for assurance that the procedures used are relevant." However, stratification has strict limits, beyond which it will have "exceeded the limitations of the data" (Rothman and Greenland, 1998). Regression methods have been shown to handle far larger numbers of variables because of the much stronger assumptions made about the nature of the relationships under study (Robins and Greenland, 1986). Indeed, unless relevant product terms are explicitly included in the multiple regression model, constancy of RR for a given independent variable across levels of all other variables is imposed.

Regression modeling of multiple exposures has traditionally involved choosing models that make strict 'dichotomized' assumptions about which variables are and are not potential confounders. The decision to limit the number of exposures that are represented simultaneously is justified if the savings in precision from a smaller model offsets any potential biases (Robins and Greenland, 1986). All the modeling approaches in the thesis involved this complex trade-off. Combining precision and bias into a single measure, the square of the variance of the estimator plus the square of the bias, results in the mean squared error (Efron, 1975), which is a common measure of estimation inaccuracy (Greenland, 2000a). This measure has been used in many studies that have advocated the gains to be had from the semi-Bayes estimation of multiple parameters (Greenland, 1992). In the present results, however, from empirical data alone, the result of the trade-off between bias and variance is not obvious, as bias cannot be estimated because the 'true' parameter values remain unknown.

As for the large numbers of chemicals that were modeled simultaneously, the common rule-of-thumb holds that the ratio of the number of parameters to be estimated to the number of observed 'outcomes' should be at least 1:10 (Harrell, Jr. et al., 1985; Peduzzi

et al., 1996). However, empirical Bayes analyses with reasonable priors can involve far more parameters than is suggested by this rule-of-thumb (Witte et al., 1994), and Hierarchical Bayesian models in general can often have more parameters than data points (Gelman et al., 1995). This is because, in multilevel models, the higher-level parameters structure some dependence among the lower-level parameters, and problems of overfitting bias are avoided. In this thesis, the semi-Bayes models were fit with only approximate fitting methods, which nevertheless still have less restrictive limitations than conventional approaches (Greenland, 1993).

An objection to the semi-Bayes models might have less to do with the Bayesian portion of the modeling and more to do with the need to estimate a single model that includes hundreds of variables. An alternative strategy might have been to change the preliminary steps of the semi-Bayes modeling, by drawing the effect estimates from separate chemical-specific regression models (Berger, 1983), such as the estimates from strategy 3. However, it should be noted that due to the presence of many high correlations among chemicals, this would be tantamount to creating a shrinkage estimator based on mutually-confounded estimates.

Semi-Bayes modeling as sensitivity analysis

The semi-Bayes models in the present thesis produced estimates for many chemicals nearly equivalent to those from the less 'data stretching' analyses, such as strategy 3. This gives some reassurance regarding the results from the conventional approach to modeling. On the other hand, in the situations where the two types of estimates in fact diverged, the semi-Bayes models might have identified an estimate from a simpler modeling approach that should be questioned. In this fashion, Greenland (2000c) has suggested that it is not necessary to make a commitment to the point estimates from the semi-Bayes model; instead, the results can be used as a form of sensitivity analysis for the results from the conventional analyses.

Although extremely implausible estimates can be a warning about problems of sparse data in finely matched or finely stratified studies (Witte et al., 1994; Greenland, 2000c), a potentially greater concern is with undetected small sample bias occurring within a reasonable range of estimated effects. Small sample bias refers to an 'away from unity'

bias in the maximum likelihood estimator, given sparse data or a relatively large ratio of parameters to data (Cordeiro and McCullagh, 1991). When this is combined with inflated estimates from small study biases, such as from misclassification and confounding (Rothman and Greenland, 1998), it can lead to estimates being spuriously statistically significant or incorrectly perceived as elevated (Greenland, 2000c). In small studies, and in large studies assessing many parameters, small sample biases can have moderately large influences, in which case bias correction is an important consideration (Cordeiro and McCullagh, 1991). The semi-Bayes approach is one example of a model that corrects for such problems.

Assumptions of exchangeability

Decisions about the choice of exchangeable categories and prior variances in strategy 8 could reasonably differ among experts. The process of justifying the information that comprises the Bayesian prior is part of the Bayesian oeuvre, but difficulties in specifying priors have always been the subject of some of the criticisms of Bayesian methods (Moore, 1997). In multiple-exposure studies, one difficulty with specifying exchangeability is that variables representing exposures should be measured on comparable scales if the parameters are to be viewed as exchangeable (Greenland, 1992). This could have been a particularly difficult problem in the present study. For instance, intensity of exposures to fumes, dusts, and liquids have no obviously common and meaningful scale on which to be measured. The issue was somewhat sidestepped by using dichotomous variables to represent exposed/unexposed histories, similar to the approaches of Thomas (1985), Greenland (1992), and de Roos (2001). A related problem has to do with how one views the exchangeability of parameters when the operational definitions of any and substantial exposure levels are on relative scales and, thus, different across chemicals. The following justification was used in the present thesis: in the absence of knowledge about the carcinogenicity of the chemicals, it is not possible to specify whether a relatively high concentration of one chemical would have a higher or lower effect on lung cancer compared to a relatively low concentration of another chemical. From this perspective, the exchangeability of their effects is a reasonable belief.

8.3 Substantive findings

The work histories of the study subjects often involved numerous exposures. Among the 3029 cancer patients used for the present analyses, 11% were not exposed to any of the 231 chemicals assessed in the thesis. At the other extreme, one individual had exposure to 96 of the chemicals at one time or other in his life. In a community-based study of occupational chemicals, the large numbers of highly correlated exposures makes attribution of cases of lung cancer to any one of these chemicals all the more difficult. Further, while it is possible to establish that some occupational exposures are associated with lung cancer in the study population, identifying which of the chemicals evaluated in the study was the carcinogen can be a challenge (more so if the true occupational carcinogen was not included in the study).

8.3.1 Priorities among the occupational chemicals

Several chemicals were flagged by the ranking and selection methods in section 7.6, and for several of these I found supporting evidence across both levels of exposure and from the various secondary analyses. As discussed in section 6.11, results worth flagging depended on the model. Statistically significant results with P-value < 0.1 were always identified. In the semi-Bayes results (strategy 8, which modeled chemical properties as well), large point estimates were also used as supporting evidence, regardless of their statistical significance. The present results did not always show confirmatory evidence of previously suspected carcinogens in this population. Several of the reasons for this are discussed in the next section, with respect to specific chemicals. For many other chemicals, the balance of evidence could be indicative of a lack of an association with lung cancer. However, 'negative' evidence from the present thesis would not be sufficiently convincing for establishing that certain chemicals do not cause lung cancer. Lung cancer is a multi-factorial disease (Samet, 1994), and any one occupational exposure would likely play but a small part. Further, the presence of several 'attenuation biases' and the uncertain mixture of concentrations (on an absolute scale) make such 'negative' evidence less convincing than the evidence supporting an etiologic role.

Table 8-1 lists all 53 chemicals selected in section 7.6. Regardless of the results in the present thesis, a few more chemicals were added to the table because they have been

evaluated by IARC. For example, beryllium compounds is listed in this table because IARC has classified it as a definite carcinogen (IARC, 1993), with suggestive evidence of lung carcinogenicity. However, its point estimates for the putative association with lung cancer were consistently below unity in most of the present analyses, and thus it was not eligible for the list of flagged chemicals in section 7.6. The table summarizes the substantive evidence from the thesis, including the results from modeling strategies 3 and 8, as well as results from the secondary analyses. To put the results in the context of previous evidence, the evaluations taken by IARC in their monograph series on occupational substances were listed alongside each chemical. IARC classifies whether chemicals are carcinogenic to humans. It does not report if chemicals are carcinogenic *for* particular cancers. The classifications reported in Table 8-1 reflect IARC's classification only if the evidence was strong or suggestive for an effect on lung cancer.

Insofar as mutual confounding among the chemicals was an issue in the main results of the simpler modeling strategies, it would equally affect all the results of the secondary analyses. As supporting evidence of the carcinogenic effects of the chemicals, then, the results on exposure characteristics and histological subtypes of lung cancer must be viewed with this potential bias in mind.

For most known occupational carcinogens, the main epidemiologic evidence of carcinogenicity has come from occupational cohort studies and animal research. The cohort studies typically involved workers with very high exposure levels (Checkoway et al., 2004). By contrast, in a community-based case-control study such as the Montreal study, exposure levels can range from very low to very high. For example, only 4 workers with a history of asbestos mining were found in the Montreal study population. The remaining asbestos-exposed workers were involved in jobs with typically low-level exposure, such as painters, carpenters, plumbers, and sheet metal workers, among others. These results, in some cases, provide evidence for whether effects manifest at such exposures levels, which is a topic of debate for even recognized carcinogens, like asbestos (LaDou et al., 2001; Siemiatycki, 2001).

Table 8-1: Summary of present results, and comparison to decisions taken by IARC monograph series

Occupational substance	*Exposure level of main results that supported an effect of the chemical* [a]		*Results from secondary analyses that supported an effect of the chemical* [b]	*IARC classification* [c]
	Strategy 3	*Strategy 8*		
127. Alkali, caustic solutions	-	Any	Small cell	NE
204. Alkanes (C5-C17)	Any, Sub	NE	-	NE
21. Aluminium alloy dust	Any	-	Window 20+ years, squamous cell	NE (1 for beryllium-aluminium alloy)
183. Aluminium compounds	Any	NE	Window 5-15 years	NE (1 for Aluminium production)
193. Arsenic compounds	-	-	-	1
5. Asbestos	Sub	-	Window 5-15 years, small cell	1
219. Benzo(a)pyrene	-	-	Squamous cell	2A
181. Beryllium compounds	-	-	-	1
24. Borates	Sub	-	Adenocarcinoma	NE
15. Brass dust	-	Sub	Window 20+ years, squamous cell	NE

[a] For the conventional (strategy 3) and semi-Bayes (strategy 8) approaches, the words ANY or SUB are listed if the estimates were earmarked by the methods of ranking and selection at those respective exposure levels, while NE (not evaluated) in the semi-Bayes column indicates the substance was not in the model of 184 chemicals. [b] Exposure characteristics are listed if the results indicated dose-response with concentration or duration, or if the exposure had its effect predominantly in one of the two time windows analyzed or with one of the three histological subtypes of lung cancer. [c] IARC evaluations are coded as 1 (carcinogenic to humans), 2A (probably carcinogenic), 2B (possibly carcinogenic), 3 (not classifiable), 4 (probably not carcinogenic), and NE (not evaluated).

Table 8-1: Summary of present results, and comparison to decisions taken by IARC monograph series

Occupational substance	*Exposure level of main results that supported an effect of the chemical* [a]		*Results from secondary analyses that supported an effect of the chemical* [b]	*IARC classification* [c]
	Strategy 3	*Strategy 8*		
195. Cadmium compounds	-	-	Window 5-15 years	1
104. Chromium fumes	Any	Any	Window 5-15 years	1 (for chromium VI compounds)
178. Chromium VI compounds	Any	NE	-	1
12. Clay dust	-	Any	Window 20+ years	NE
97. Coal gas	-	-	-	1
163. Coal tar and pitch	-	-	-	1
191. Copper compounds	Sub	NE	-	NE
35. Copper dust	Sub	Sub	Concentration	NE
108. Copper fumes	Any, Sub	Any	Window 5-15 years	NE
23. Cosmetic talc	-	Any	-	3
6. Crystalline silica	Any, Sub	-	Duration	1

[a] For the conventional (strategy 3) and semi-Bayes (strategy 8) approaches, the words ANY or SUB are listed if the estimates were earmarked by the methods of ranking and selection at those respective exposure levels, while NE (not evaluated) in the semi-Bayes column indicates the substance was not in the model of 184 chemicals. [b] Exposure characteristics are listed if the results indicated dose-response with concentration or duration, or if the exposure had its effect predominantly in one of the two time windows analyzed or with one of the three histological subtypes of lung cancer. [c] IARC evaluations are coded as 1 (carcinogenic to humans), 2A (probably carcinogenic), 2B (possibly carcinogenic), 3 (not classifiable), 4 (probably not carcinogenic), and NE (not evaluated).

Table 8-1: Summary of present results, and comparison to decisions taken by IARC monograph series

Occupational substance	*Exposure level of main results that supported an effect of the chemical* [a]		*Results from secondary analyses that supported an effect of the chemical* [b]	*IARC classification* [c]
	Strategy 3	*Strategy 8*		
161. Cutting fluids	Any	-	Window 5-15 years	NE
170. Cutting fluids pre 1955	-	NE	-	1 (for mineral oils)
171. Cutting fluids post 1955	Any	NE	Window 5-15 years	NE
117. Diesel engine emissions	-	-	-	2A
3. Excavation dust	Sub	-	Duration, window 5-15 years, small cell	NE
177. Fluorides	Any	NE	Window 5-15 years	3
98. Gas welding fumes	Any	-	Concentration	2B
115. Gasoline engine emissions	-	-	-	2B
8. Glass dust	Any	-	Window 5-15 years, squamous cell	2A (for manufacturing)
9. Glass fibres	-	-	-	3, 2B
213. Glycol ethers	Sub	-	Concentration	3

[a] For the conventional (strategy 3) and semi-Bayes (strategy 8) approaches, the words ANY or SUB are listed if the estimates were earmarked by the methods of ranking and selection at those respective exposure levels, while NE (not evaluated) in the semi-Bayes column indicates the substance was not in the model of 184 chemicals. [b] Exposure characteristics are listed if the results indicated dose-response with concentration or duration, or if the exposure had its effect predominantly in one of the two time windows analyzed or with one of the three histological subtypes of lung cancer. [c] IARC evaluations are coded as 1 (carcinogenic to humans), 2A (probably carcinogenic), 2B (possibly carcinogenic), 3 (not classifiable), 4 (probably not carcinogenic), and NE (not evaluated).

Table 8-1: Summary of present results, and comparison to decisions taken by IARC monograph series

Occupational substance	*Exposure level of main results that supported an effect of the chemical* [a]		*Results from secondary analyses that supported an effect of the chemical* [b]	*IARC classification* [c]
	Strategy 3	*Strategy 8*		
158. Heating oil	Sub	-	Duration	3 (light distillates)
85. Hydrogen fluoride	Any	-	Window 5-15 years	3
174. Inks	Any	-	-	3
126. Inorganic acid solutions	-	-	-	1 (for acid mists containing sulphuric acid)
156. Kerosene	Any, Sub	Sub	Concentration, window 5-15 years	3
40. Lead chromate	-	-	-	1
201. Lead compounds	-	NE	-	2A (for inorganic lead) and 3 (for organic lead)
182. Magnesium compounds	Any	NE	Concentration, squamous cell	NE
187. Manganese compounds	Sub	NE	Duration, window 5-15 years	NE
105. Manganese fumes	Sub	Sub	Window 5-15 years	NE

[a] For the conventional (strategy 3) and semi-Bayes (strategy 8) approaches, the words ANY or SUB are listed if the estimates were earmarked by the methods of ranking and selection at those respective exposure levels, while NE (not evaluated) in the semi-Bayes column indicates the substance was not in the model of 184 chemicals. [b] Exposure characteristics are listed if the results indicated dose-response with concentration or duration, or if the exposure had its effect predominantly in one of the two time windows analyzed or with one of the three histological subtypes of lung cancer. [c] IARC evaluations are coded as 1 (carcinogenic to humans), 2A (probably carcinogenic), 2B (possibly carcinogenic), 3 (not classifiable), 4 (probably not carcinogenic), and NE (not evaluated).

Table 8-1: Summary of present results, and comparison to decisions taken by IARC monograph series

Occupational substance	*Exposure level of main results that supported an effect of the chemical* [a]		*Results from secondary analyses that supported an effect of the chemical* [b]	*IARC classification* [c]
	Strategy 3	*Strategy 8*		
101. Metal oxide fumes	Sub	-	Concentration	NE (Evaluated individually)
4. Metallic dust	Any, Sub	NE	Concentration, window 5-15 years	NE (Evaluated individually)
91. Methane	Sub	Sub	-	3
90. Natural gas	Sub	NE	Concentration	NE (evaluated individually)
53. Natural rubber	-	Any	-	NE (1 for rubber industry)
190. Nickel compounds	Any	NE	Window 5-15 years	1 (2B for metallic nickel)
107. Nickel fumes	Any	-	Window 5-15 years	1
83. Nitrogen oxides	Any, Sub	Any	-	NE
214. PAHs	-	-	Concentration, squamous cell	Depends on particular PAH
64. Poly-acrylates	-	Any	-	3 (for polymethyl methacrylate)
122. Propane engine emissions	Any	-	Squamous cell	NE

[a] For the conventional (strategy 3) and semi-Bayes (strategy 8) approaches, the words ANY or SUB are listed if the estimates were earmarked by the methods of ranking and selection at those respective exposure levels, while NE (not evaluated) in the semi-Bayes column indicates the substance was not in the model of 184 chemicals. [b] Exposure characteristics are listed if the results indicated dose-response with concentration or duration, or if the exposure had its effect predominantly in one of the two time windows analyzed or with one of the three histological subtypes of lung cancer. [c] IARC evaluations are coded as 1 (carcinogenic to humans), 2A (probably carcinogenic), 2B (possibly carcinogenic), 3 (not classifiable), 4 (probably not carcinogenic), and NE (not evaluated).

Table 8-1: Summary of present results, and comparison to decisions taken by IARC monograph series

Occupational substance	*Exposure level of main results that supported an effect of the chemical* [a]		*Results from secondary analyses that supported an effect of the chemical* [b]	*IARC classification* [c]
	Strategy 3	*Strategy 8*		
76. Soot	-	-	Concentration	1
196. Tin compounds	Any	NE	-	NE
68. Urea-formaldehyde	-	Any	Window 5-15 years, squamous cell	NE
192. Zinc compounds	Sub	NE	Concentration	NE
36. Zinc dust	Sub	-	Concentration	NE
109. Zinc fumes	Sub	-	Concentration	NE

[a] For the conventional (strategy 3) and semi-Bayes (strategy 8) approaches, the words ANY or SUB are listed if the estimates were earmarked by the methods of ranking and selection at those respective exposure levels, while NE (not evaluated) in the semi-Bayes column indicates the substance was not in the model of 184 chemicals. [b] Exposure characteristics are listed if the results indicated dose-response with concentration or duration, or if the exposure had its effect predominantly in one of the two time windows analyzed or with one of the three histological subtypes of lung cancer. [c] IARC evaluations are coded as 1 (carcinogenic to humans), 2A (probably carcinogenic), 2B (possibly carcinogenic), 3 (not classifiable), 4 (probably not carcinogenic), and NE (not evaluated).

8.3.2 Comments on selected chemicals

The following comments refer to the findings of several of the chemicals evaluated in the thesis, and it places these results in the context of previously published evidence. For space considerations, not all the chemicals are discussed.

Aluminium alloy dust. The present results are consistent with previous evidence of lung cancer risks in the aeronautical industry, occasionally attributed to a beryllium-aluminium alloy (IARC, 1993). Although aluminium compounds per se have not been evaluated by IARC, aluminium production workers have been identified as having higher risk of lung cancer (IARC, 1987a). This has been attributed to exposure to coal tars (Armstrong and Theriault, 1996) and polycyclic aromatic hydrocarbons (Armstrong et al., 1994).

Arsenic compounds. Arsenic compounds were not flagged as possibly carcinogenic in the present results, possibly due to the diverse forms of arsenic chemicals that were included in this group. Increases in the risk of lung cancer have been observed in epidemiologic studies of copper smelter workers, who have heavy exposures to arsenic (Lee-Feldstein, 1986), as well as decreases in risk following cessation of exposure among workers at a copper smelter (Enterline and Marsh, 1980). One study indicated that arsenic may have a unique concave form to its exposure-response curve (Enterline et al., 1987a), which could also explain the low estimates observed here.

Asbestos. Results for asbestos at any level of exposure were consistent with the lack of an effect on lung cancer. While different modeling strategies resulted in different inferences, the estimates of at least one model at substantial levels of exposure were suggestive of a small effect. Secondary results indicated an effect on squamous cell tumours. In those few previous studies that evaluated asbestos and histological types of lung cancer, there was no evidence of a specificity of effect on particular cell types (Churg, 1994). A limitation of the present results is that there was no distinction of the effects of amphibole- and chrysotile-type asbestos exposures, for which there is considerable controversy as to their relative roles in lung cancer (Henderson et al., 2004). Some evidence suggests that the effect of asbestos occurs only 15 or 20 years (and possibly up to 35 years) after initial exposure (Selikoff et al., 1979). The present secondary results suggest that the relevant exposure was primarily in the 5 to 15 years prior to diagnosis.

The majority of the exposure to asbestos in the Montreal study was of the chrysotile form, though amphibole exposures were also involved. A recent attempt at pooling studies of only chrysotile exposures resulted in a meta-analysis summary SMR for lung cancer of 2.4 (Li et al., 2004). These findings were primarily among miners, asbestos producers, and textile workers, and exposures in those industries are known to be orders of magnitude higher than exposure levels in the bulk of the industries in which Montreal workers received their asbestos exposure. There remains controversy as to the effects of low-level exposures and as to whether workplaces are sufficiently regulated (LaDou et al., 2001; Camus, 2001; Siemiatycki, 2001). A recent review (Laden et al., 2004) and a recent meta-analysis (Goodman et al., 2004) of motor vehicle mechanics, who would be expected to have low exposures to asbestos as a result of automobile brake repair, concluded that evidence did not support an increased risk of lung cancer.

Cadmium compounds. Evaluation of cadmium has been difficult due to the frequent concomitant exposure with other recognized lung carcinogens (IARC, 1993). Aside from a slightly elevated estimate for exposure in the 5 to 15 years prior to diagnosis, the present results offer little to support the hypothesis of a causal relation. Although cadmium has been categorized as a definite carcinogen by IARC, results from occupational studies have been inconsistent (World Health Organization, 1992) and at least one recent study has not supported cadmium's lung carcinogenicity (Sorahan and Esmen, 2004). Recent recommendations have downgraded the level of certainty that cadmium is carcinogenic to humans (Jarup, 2003).

Crystalline silica. Silica has been associated with several serious illnesses, including lung cancer (Steenland, 2005). In the present results, the estimated effect of silica disappeared when adjusted for the effects of other chemicals. This may have been related to an over-adjustment of silica's estimate in the semi-Bayes models because of the inclusion of other chemicals chiefly composed of silica, like excavation dust, Portland cement, and brick dust. All the modeling strategies showed limited evidence of an increasing exposure-response curve, including the results of analyzing duration in the secondary analyses. Several studies have identified dose-surrogates for silica exposure, like duration, that are related to the risk of lung cancer (IARC, 1997). On the other hand, the relationship with duration has not been consistently observed (Hughes et al., 2001).

Cutting fluids. Cutting fluids used before 1955 were composed largely of mineral oils. Treated mineral oils have been evaluated by IARC as definitely carcinogenic for some forms of cancer (IARC, 1987a), but evidence for lung carcinogenicity has been less consistent. Higher lung cancer mortality has been seen in metalworkers (Acquavella et al., 1993; Kazerouni et al., 2000), and effects are often attributed to the release of PAHs. The present results do not support an effect of early exposure to cutting fluids on lung cancer. Oddly, several elevated estimates were noted with exposures to cutting fluids formulated after 1955, which purportedly release less PAHs and tend to be comprised of emulsified mineral oils and synthetic fluids.

Diesel engine emissions. Various approaches were attempted to assess the effect of diesel exhaust, such as whether or not to adjust for the effects of other whole engine exhausts and particular engine emissions. In all formulations of the regression models, the point estimates for diesel exhaust did not change appreciably and were always consistent with little or no effect on the risk of lung cancer. Many of the emissions of diesel are themselves suspected carcinogens, some being common to cigarette smoke. The evaluation of diesel exhaust by IARC as a probable carcinogen (IARC, 1989) was not maintained in the later evaluation of the Environmental Protection Agency (Hughes et al., 2001), mostly due to the perceived weaknesses of many of the epidemiologic studies. Further, with current stricter emission control standards, past studies of typical diesel exposure may not be appropriate guides to the effects of current exposures (Bunn, III et al., 2004).

Glass dusts and fibres. Manufacturers of art glassworks have been evaluated by IARC as probably carcinogenic (IARC, 1987a), though a recent update to a large cohort found that the previously elevated risks (Wingren and Englander, 1990) were no long evident (Wingren, 2004), possibly due to improved workplace hygiene. Glass dust exposure was flagged in the results of the conventional models, though this elevated estimate disappeared upon adjustment for the effects of other occupational chemicals. All the present results for the exposure to glass fibres were consistent with a lack of an effect on lung cancer. The recent evaluations of man-made vitreous fibres by IARC (1988; 2002) were mainly based on two epidemiologic studies (Boffetta et al., 1999; Marsh et al., 2001a; Marsh et al., 2001b). Exposures to insulation glass wool and continuous glass

filament remained unclassifiable as to carcinogenicity, while special-purpose glass fibres were classified as possibly carcinogenic.

Lead compounds. The present results on lead compounds, including lead chromate, lead fumes, and lead oxides showed consistent evidence of a lack of an effect on lung cancer. Although IARC has evaluated inorganic lead compounds as possibly carcinogenic (IARC, 1987a), a recent meta-analysis suggests only a possibly weak association with lung cancer (Steenland and Boffetta, 2000a). Concomitant exposures, such as to arsenic compounds, have presented difficulties with the interpretation of epidemiologic studies.

Nickel and chromium compounds. The present results showed only an inconsistent association of nickel compounds with lung cancer. Separating the effect of nickel fumes from chromium fumes was, for the most part, not possible due to the high correlation between the two exposures, and there was insufficient exposure at substantial levels to estimate the effect of nickel fumes in this study population. Nevertheless, in a model adjusting for non-occupational characteristics and seven lung carcinogens, which included the general category of chromium VI compounds, a statistically elevated point estimate resulted for nickel fumes. This estimated effect disappeared upon inclusion of the rest of the occupational chemicals, but the attenuation was driven by the presence of chromium fumes in the model. Distinguishing these chemicals in the semi-Bayes prior (strategy 8), resulted in the estimate of nickel fumes remaining close to unity while that of chromium fumes was slightly elevated (though not statistically). Chromium fumes and hexavalent chromium in general were both flagged in several of the present results, including an elevated estimate for chromium fumes in the 5 to 15 years prior to diagnosis. The majority of nickel exposure in the Montreal study consisted of metallic nickel, which IARC has evaluated as only possibly carcinogenic to humans (IARC, 1990). Studies of lung cancer and nickel have mostly identified water-soluble nickel as carcinogenic (Grimsrud et al., 2002). A recent study, however, supported the excess risk of lung cancer among nickel refinery workers, even after eliminating the effects of other potential carcinogenic chemicals (Grimsrud et al., 2005).

Nitrogen oxides. Nitrogen oxides (NO) have not been evaluated by IARC, and the World Health Organization (1997) concluded that insufficient evidence existed to evaluate their

carcinogenic potency. An observed excess risk of lung cancer has been linked to environmental air pollution, particularly to levels of NO (Nafstad et al., 2003). Although that study used NO measurement as a surrogate for air pollution and exposure to other possible carcinogens, at least some evidence suggests that NO derivatives can play a part in lung carcinogenesis (Masri et al., 2005). A weak association has also been observed with lung cancer mortality and fertilizer workers (Bulbulyan et al., 1996). The present results for NO were consistent across models, and the magnitude of the effect on lung cancer was one of the strongest among the chemicals assessed.

Polycyclic aromatic hydrocarbons. PAHs are ubiquitous exposures common to air pollution, cooking fumes, and many occupational sources. Benzo(a)pyrene, in particular, has been evaluated as definitely carcinogenic to humans (IARC, 1983), but epidemiologic evidence in general has not been sufficient to evaluate the carcinogenic effects of PAHs (Boffetta et al., 1997). The present results from the main modeling strategies did not support elevated risks in respect to PAH exposures, however caution is necessary when interpreting these particular results because of the strong assumptions made in all the models. Secondary results, however, flagged effects in relation specifically to squamous cell tumours, and an exposure-response trend of concentration was evident when PAHs were considered as a whole group. In the sensitivity analyses, weak evidence pointed to elevated risk of PAHs originating from coal. The exposure to PAHs from 'other sources' (id 215), referring to PAHs from food, plastics, and paints, also had an elevated point estimate. Interpreting these results is complicated by the complex relationships among PAHs and other chemicals that adsorb them.

Soot. The evaluation of soot and lung cancer is based on only a few epidemiologic studies (IARC, 1987a). Among chimney sweeps, excess lung cancer mortality has been observed (Hogstedt et al., 1982), with some limited evidence for a specificity for small cell lung carcinomas (Evanoff et al., 1993). The present results were inconsistent, with only the secondary results for a trend with concentration being statistically significant.

Several recognized lung carcinogens were not flagged in the current results, including coal gas, coal tar and pitch, beryllium compounds, and lead chromate. Several reasons could be postulated for why these chemicals did not result in elevated estimates in this

study population. One reason is that they may have occurred predominantly at low exposure levels. A second reason is that each occurred with low prevalence, and the associated estimates were thus imprecise.

8.4 Summary

This thesis focused on the effects of 231 occupational chemicals on the risk of lung cancer. The International Agency for Research on Cancer has evaluated many of these substances, and some attempt was made to place the present results in the context of this 'previous evidence.' However, the nature of the list of 231 chemicals, which includes many overlapping chemical groups and complex mixtures, posed some difficulty with this task. The following can be taken as a crude guide to the overall results. IARC has classified about fourteen of the 231 chemicals as definite carcinogens, with strong evidence for lung carcinogenicity, and three as probable carcinogens. Among these, the following eight chemicals were found to have at least some supporting evidence in the present thesis: asbestos, benzo(a)pyrene, cadmium compounds, chromium fumes, chromium VI compounds, crystalline silica, nickel fumes, and soot. For six of the chemicals previously suspected of being carcinogenic for lung, namely arsenic compounds, beryllium compounds, coal gas, coal tar and pitch, diesel exhaust, and lead chromate, there was little evidence of an association in the Montreal study population, though low prevalence may have contributed to this. Of the other chemicals that have not been evaluated by IARC, the current results offer suggestive evidence of an increased risk of lung cancer due to these exposures: caustic alkali solutions (such as sodium and potassium hydroxides), aluminium compounds, borates, brass dust, clay dust, copper compounds, copper dust, copper fumes, various cutting fluids, excavation dust, magnesium compounds, manganese compounds, nitrogen oxides, propane exhaust, tin compounds, urea-formaldehyde, and zinc compounds. Finally, for many of the remaining chemicals, the balance of evidence from the results would be indicative of a lack of an association with lung cancer. On the whole, the results from this study provided some unique evidence on the risk of lung cancer in relation to a large range of occupational chemicals, for which little is known about their carcinogenic effects.

A key feature of this thesis was the adjustment for mutual confounding among the chemical exposures. Eight modeling strategies were used to assess the effects of the chemicals, from simple models that assumed no confounding among the occupational exposures, to complex models adjusting for all effects simultaneously in a Bayesian framework. The number of high correlations among the chemicals would have expectedly led to overestimation of many effects if the correlations between irrelevant chemicals and genuine lung carcinogens had not been accounted for. Indeed, the estimated effects of many chemicals appeared to be inflated in the simpler modeling strategies.

The use of a single logistic regression model avoided what are often ad hoc decisions and methods for choosing which chemicals will be included as confounders and which chemicals will not be. However, this model could not have been efficiently fit without applying it within a semi-Bayes framework, which improves estimation by shrinking implausible estimates to more reasonable values in the face of biases caused by sparse data. Furthermore, the semi-Bayes model allowed for the inclusion of near-collinear variables, a situation that is known to introduce a type of confounding that typically cannot be resolved in conventional analyses.

The semi-Bayes model incorporated expert opinions at the stage of analysis, by grouping sets of chemicals by their shared chemical properties. In principle, this extra information should have resulted in more accurate estimates for all the chemicals analyzed. In fact, the inclusion of chemical properties did not materially influence the results over a simpler form of the semi-Bayes model. This perhaps reflects the uncertain nature of the current state of knowledge about the relevant characteristics involved in occupational carcinogenesis.

In conclusion, this doctoral dissertation provides a large body of evidence towards a better understanding of the risk of lung cancer, as it pertains to the many chemical exposures that are widespread in the workplace. The application of semi-Bayes modeling offered several benefits over more conventional approaches, but any gains predicated on theory should be tempered with an understanding of some of the limitations of modeling hundreds of chemical effects in a single model.

8.5 Suggestions for future research

A natural extension of the present work would be to follow-up on several of the chemicals that the results of the thesis identified as possibly carcinogenic for lung cancer. Since the results from the Montreal study represent the only evidence available for many of the chemicals, replications would be necessary to distinguish the 'false positives' from the 'true positives.' Even among those chemicals previously suspected of being carcinogenic, past assessments were occasionally based on limited epidemiologic evidence, and so follow-up of well-known chemicals is necessary.

The semi-Bayes model is an attractive approach to analyses of multiple exposures that has been available for decades and yet has not been widely used. More work is necessary to understand the tradeoffs between possible drawbacks and gains from undertaking this more complex analytic approach. In particular, a simulation study would be necessary to address some of the questions that were raised (and unanswered) by the present thesis. The present work involved only a descriptive comparison of the results from the different modeling strategies, but for the next step, a valuable contribution would involve an evaluation of the relative accuracies of the semi-Bayes model and the approach involving a separate regression model for each chemical. Surprisingly, such an analysis was missing from comparisons performed in previous simulation studies (Greenland, 1993; Witte and Greenland, 1996). Furthermore, these studies have tended to simulate data involving smaller samples and fewer parameters than the Montreal study, and so those results may not have been applicable to the present application of semi-Bayes modeling.

With so many chemicals assessed simultaneously, improvements in the semi-Bayes model might be possible. Further estimation accuracy might be gained by adding a third level to the hierarchy, involving higher parameters that would describe the chemical properties themselves. Exploration of such 'Bayesian empirical Bayes' models might be warranted (Greenland, 2000a; Robert, 2001). Other applications of semi-Bayes models in occupational cancer research are also possible. Analyses at the level of occupational groups or industries have historically been a fruitful source of information about occupational carcinogens, and an application of semi-Bayes models to these analyses might also serve to improve estimation. The second-level model, which imposes the

exchangeability assumptions, could be based on whether the occupations or industries share a similar profile of known carcinogens. There are likely many other worthwhile applications of semi-Bayes models in occupational cancer research and epidemiologic research, in general.

9 References

Acquavella J, Leet T, and Johnson G (1993) Occupational experience and mortality among a cohort of metal components manufacturing workers. *Epidemiology* **4**: 428-434.

Agricola G. *De Re Metallica, Translated from the First Latin Edition of 1556, by Hoover, H.C. and Hoover, L.H.,* New York: Dover Publications, Inc., tr1950.

Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**: 716-723.

Alavanja MC, Brownson RC, Boice JD, Jr., and Hock E (1992) Preexisting lung disease and lung cancer among nonsmoking women. *American Journal of Epidemiology* **136**: 623-632.

Alberg AJ and Samet JM (2003) Epidemiology of lung cancer. *Chest* **123**: 21s-49s.

Amandus H and Costello J (1991) Silicosis and lung cancer in U.S. metal miners. *Archives of Environmental Health* **46**: 82-89.

American Medical Association. *American Medical Association, International Classification of Diseases, Clinical Modification, 9th Edition,* Chicago: 1998.

Armstrong B, Tremblay C, Baris D, and Theriault G (1994) Lung cancer mortality and polynuclear aromatic hydrocarbons: a case-cohort study of aluminum production workers in Arvida, Quebec, Canada. *American Journal of Epidemiology* **139**: 250-262.

Armstrong BG (1998) Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occupational & Environmental Medicine* **55**: 651-656.

Armstrong BG and Theriault G (1996) Compensating lung cancer patients occupationally exposed to coal tar pitch volatiles. *Occupational & Environmental Medicine* **53**: 160-167.

Axelson O and Sjoberg A (1979) Cancer incidence and exposure to iron oxide dust. *J Occup Med* **21**: 419-422.

Bandera EV, Freudenheim JL, and Vena JE (2001) Alcohol consumption and lung cancer: a review of the epidemiologic evidence. *Cancer Epidemiology Biomarkers and Prevention* **10**: 813-821.

Barbone F, Bovenzi M, Cavallieri F, and Stanta G (1997) Cigarette smoking and histologic type of lung cancer in men. *Chest* **112**: 1474-1479.

Beadsmoore CJ and Screaton NJ (2003) Classification, staging and prognosis of lung cancer. *Eur J Radiol* **45**: 8-17.

BEIR IV (1988) Committee on Biological Effects of Ionizing Radiation: Health Risks of Radon and Other Internally Deposited Alpha-Emitters. United States National Academy of Science, National Research Council, Washington, D.C.: National Academy Press.

Berenblum I. *Cancer Research Today,* London: Pergamon Press, 1967.

Berger J (1983) Parametric empirical Bayes: theory and application: Comment. *Journal of the American Statistical Association* **78**: 55-57.

Berry G, Newhouse ML, and Antonis P (1985) Combined effect of asbestos and smoking on mortality from lung cancer and mesothelioma in factory workers. *British Journal of Industrial Medicine* **42**: 12-18.

Binks K, Doll R, Gillies M, Holroyd C, Jones SR, McGeoghegan D, Scott L, Wakeford R, and Walker P (2005) Mortality experience of male workers at a UK tin smelter. *Occup Med (Lond)* **55**: 215-226.

Bjelke E (1975) Dietary vitamin A and human lung cancer. *International Journal of Cancer* **15**: 561-565.

Blair A, Stewart P, O'Berg M, Gaffey W, Walrath J, Ward J, Bales R, Kaplan S, and Cubit D (1986) Mortality among industrial workers exposed to formaldehyde. *J Natl Cancer Inst* **76**: 1071-1084.

Blot WJ and Fraumeni JF, Jr. Arsenic and Lung Cancer. In: Samet JM, ed., *The Epidemiology of Lung Cancer.* New York: Marcell Dekker, 1994.

Blot WJ and Fraumeni JF, Jr. Chapter 31: Cancers of the lung and pleura. In: Schottenfeld D and Fraumeni JF, Jr., eds., *Cancer Epidemiology and Prevention.* New York: Oxford University Press, 1996.

Blot WJ, Harrington JM, Toledo A, Hoover R, Heath CW, Jr., and Fraumeni JF, Jr. (1978) Lung cancer after employment in shipyards during World War II. *New England Journal of Medicine* **299**: 620-624.

Boffetta P. Methodological aspects of the epidemiologic association between cadmium and cancer in humans. In: Nordberg G, Herber R et al, eds., *Cadmium in the Human Environment: Toxicity and Carcinogenicity*. Lyon, France: IARC, 1992.

Boffetta P (2004) Epidemiology of environmental and occupational cancer. *Oncogene* **23**: 6392-6403.

Boffetta P, Andersen A, Hansen J, Olsen JH, Plato N, Teppo L, Westerholm P, and Saracci R (1999) Cancer incidence among European man-made vitreous fiber production workers. *Scand J Work Environ Health* **25**: 222-226.

Boffetta P, Harris RE, and Wynder EL (1990) Case-control study on occupational exposure to diesel exhaust and lung cancer risk. *American Journal of Industrial Medicine* **17**: 577-591.

Boffetta P, Jourenkova N, and Gustavsson P (1997) Cancer risk from occupational and environmental exposure to polycyclic aromatic hydrocarbons. *Cancer Causes Control* **8**: 444-472.

Boffetta P, Kogevinas M, Simonato L, Wilbourn J, and Saracci R (1995) Current Perspectives on Occupational Cancer Risks. *International Journal Occupational and Environmental Health* **1**: 315-325.

Boffetta P and Nyberg F (2003) Contribution of environmental factors to cancer risk. *British Medical Bulletin* **68**: 71-94.

Boffetta P, Stellman SD, and Garfinkel L (1988) Diesel exhaust exposure and mortality among males in the American Cancer Society prospective study. *American Journal of Industrial Medicine* **14**: 403-415.

Bridge JC (1933) Annual report of the chief inspector for the year 1932. London: Her Majesty's Stationary Office.

Brown CC and Chu KC (1983) Implications of the multistage theory of carcinogenesis applied to occupational arsenic exposure. *J Natl Cancer Inst* **70**: 455-463.

Browne K (1986) Is asbestos or asbestosis the cause of the increased risk of lung cancer in asbestos workers? *British Journal of Industrial Medicine* **43**: 145-149.

Brownson RC and Alavanja MC (2000) Previous lung disease and lung cancer risk among women (United States). *Cancer Causes Control* **11**: 853-858.

Buffler PA, Wood S, Eifler C, Suarez L, and Kilian DJ (1979) Mortality experience of workers in a vinyl chloride monomer production plant. *Journal of Occupational Medicine* **21**: 195-203.

Bulbulyan MA, Jourenkova NJ, Boffetta P, Astashevsky SV, Mukeria AF, and Zaridze DG (1996) Mortality in a cohort of Russian fertilizer workers. *Scand J Work Environ Health* **22**: 27-33.

Bunn WB, III, Hesterberg TW, Valberg PA, Slavin TJ, Hart G, and Lapin CA (2004) A reevaluation of the literature regarding the health assessment of diesel engine exhaust. *Inhal Toxicol* **16**: 889-900.

Burns DM. Chapter 2: Tobacco Smoking. In: Samet JM, ed., *Epidemiology of Lung Cancer*. New York: Marcel Dekker, Inc., 1994.

Byers T. Chapter 13: Diet as a Factor in the Etiology and Prevention of Lung Cancer. In: Samet JM, ed., *Epidemiology of Lung Cancer*. New York: Marcel Dekker, Inc., 1994.

Camus M (2001) A ban on asbestos must be based on a comparative risk assessment. *Canadian Medical Association Journal* **164**: 491-494.

Casella G (1985) An introduction to empirical Bayes data analysis. *The American Statistician* **39**: 83-87.

Chavance M, Dellatolas G, and Lellouch J (1992) Correlated nondifferential misclassifications of disease and exposure: application to a cross-sectional study of the relation between handedness and immune disorders. *International Journal of Epidemiology* **21**: 537-546.

Checkoway H, Heyer NJ, Seixas NS, Welp EA, Demers PA, Hughes JM, and Weill H (1997) Dose-response associations of silica with nonmalignant respiratory disease and lung cancer mortality in the diatomaceous earth industry. *American Journal of Epidemiology* **145**: 680-688.

Checkoway H, Pearce N, and Kriebel D. *Occupational Epidemiology, 2nd Edition,* Toronto: Oxford University Press, 2004.

Chen JL, Fayerweather WE, and Pell S (1988a) Cancer incidence of workers exposed to dimethylformamide and/or acrylonitrile. *Journal of Occupational Medicine* **30**: 813-818.

Chen JL, Fayerweather WE, and Pell S (1988b) Mortality study of workers exposed to dimethylformamide and/or acrylonitrile. *Journal of Occupational Medicine* **30**: 819-821.

Churg A. Chapter 16: Lung Cancer Cell Type and Occupational Exposure. In: Samet JM, ed., *Epidemiology of Lung Cancer*. New York: Marcel Dekker, Inc., 1994.

Clarke RH and Southwood TR (1989) Risks from ionizing radiation. *Nature* **338**: 197-198.

Cole P (1979) The evolving case-control study. *Journal of Chronic Diseases* **32**: 15-27.

Collingwood KW, Pasternack BS, and Shore RE (1987) An industry-wide study of respiratory cancer in chemical workers exposed to chloromethyl ethers. *J Natl Cancer Inst* **78**: 1127-1136.

Collins JJ, Page LC, Caporossi JC, Utidjian HM, and Lucas LJ (1989) Mortality patterns among employees exposed to acrylonitrile. *Journal of Occupational Medicine* **31**: 368-371.

Cordeiro GM and McCullagh P (1991) Bias correction in generalized linear models. *Journal of the Royal Statistical Society, Series B* **53**: 629-643.

Costello J and Graham WG (1988) Vermont granite workers' mortality study. *American Journal of Industrial Medicine* **13**: 483-497.

Darby S, Hill D, and Doll R (2001) Radon: a likely carcinogen at all exposures. *Annals of Oncology* **12**: 1341-1351.

Darby SC and Pike MC (1988) Lung cancer and passive smoking: predicted effects from a mathematical model for cigarette smoking and lung cancer. *British Journal Of Cancer* **58**: 825-831.

Davies JM (1984) Lung cancer mortality among workers making lead chromate and zinc chromate pigments at three English factories. *British Journal of Industrial Medicine* **41**: 158-169.

De Roos AJ, Poole C, Teschke K, and Olshan AF (2001) An application of hierarchical regression in the investigation of multiple paternal occupational exposures and

neuroblastoma in offspring. *American Journal of Industrial Medicine* **39**: 477-486.

De Roos AJ, Rothman N, Inskip PD, Linet MS, Shapiro WR, Selker RG, Fine HA, Black PM, Pittman GS, and Bell DA (2003) Genetic polymorphisms in GSTM1, -P1, -T1, and CYP2E1 and the risk of adult brain tumors. *Cancer Epidemiology Biomarkers and Prevention* **12**: 14-22.

DHHS (1989) Department of Health and Human Services: Reducing the health consequences of smoking: 25 years of progress; a report of the Surgeon General. Washington, DC: US Government Printing Office.

DHHS (1990) Department of Health and Human Services: The health benefits of smoking cessation. A report of the Surgeon General. Washington, D.C.: US Government Printing Office.

Doll R (1955) Mortality from lung cancer in asbestos workers. *British Journal of Industrial Medicine* **12**: 81-86.

Doll R (1959) Occupational lung cancer: a review. *British Journal of Industrial Medicine* **16**: 181-190.

Doll R (1975) Pott and the pospects for prevention. *British Journal Of Cancer* **32**: 263-272.

Doll R. Chapter 1: Introduction and overview. In: Samet JM, ed., *Epidemiology of Lung Cancer*. New York: Marcel Dekker, Inc., 1994.

Doll R and Hill AB (1950) Smoking and carcinoma of the lung; preliminary report. *British Medical Journal* **2**: 739-748.

Doll R, Mathews JD, and Morgan LG (1977) Cancers of the lung and nasal sinuses in nickel workers: a reassessment of the period of risk. *British Journal of Industrial Medicine* **34**: 102-105.

Doll R and Peto J (1985) Asbestos. Effects on health of exposure to asbestos. London: Health and Safety Commission, Her Majesty's Stationary Office.

Doll R and Peto R (1978) Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. *J Epidemiol Community Health* **32**: 303-313.

Doll R and Peto R (1981) The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst* **66**: 1191-1308.

Doll R, Peto R, Wheatley K, Gray R, and Sutherland I (1994) Mortality in relation to smoking: 40 years' observations on male British doctors. *British Medical Journal* **309**: 901-911.

Doll R, Vessey MP, Beasely RWR, Buckley AR, Fear EC, Fisher REW, Gammon EJ, Gunn W, Hughes GO, Lee K, and Norman-Smith B (1972) Mortality of gas workes - final report of a prospective study. *British Journal of Industrial Medicine* **29**: 394-406.

Economou P, Samet JM, and Lechner JF. Familial Genetic Factors in the Pathogenesis of Lung Cancer. In: Samet JM, ed., *Epidemiology of Lung Cancer.* New York: Marcel Dekker, Inc., 1994.

Efron B (1975) Biased versus unbiased estimation. *Advanced Mathematics* **16**: 259-277.

Efron B and Morris CN (1973) Stein's estimation rule and its competitors -- an empirical Bayes approach. *Journal of the American Statistical Association* **68**: 117-130.

Emmelin A, Nystrom L, and Wall S (1993) Diesel exhaust exposure and smoking: a case-referent study of lung cancer among Swedish dock workers. *Epidemiology* **4**: 237-244.

Enterline PE (1974) Respiratory cancer among chromate workers. *Journal of Occupational Medicine* **16**: 523-526.

Enterline PE, Henderson VL, and Marsh GM (1987a) Exposure to arsenic and respiratory cancer. A reanalysis. *American Journal of Epidemiology* **125**: 929-938.

Enterline PE and Marsh GM (1980) Mortality studies of smelter workers. *Am J Ind Med* **1**: 251-259.

Enterline PE, Marsh GM, Esmen NA, Henderson VL, Callahan CM, and Paik M (1987b) Some effects of cigarette smoking, arsenic, and SO2 on mortality among US copper smelter workers. *Journal of Occupational Medicine* **29**: 831-838.

Evanoff BA, Gustavsson P, and Hogstedt C (1993) Mortality and incidence of cancer in a cohort of Swedish chimney sweeps: an extended follow up study. *Br J Ind Med* **50**: 450-459.

Farr W. *Vital Statistics: A Memorial Volume of Selections from the Reports and Writings of William Farr, 1885,* Reprint, Metuchen, N.J.: Scarecrow Press, 1975.

Fraumeni JF, Jr. and Blot WJ. Lung and Pleura. In: Schottenfeld D and Fraumeni JF, Jr., eds., *Cancer Epidemiology and Prevention.* Philadelphia: Saunders, 1982.

Freedman L and Pee D (1989) Return to a note on screening regression equations. *American Statistician* **43**: 279-282.

Frentzel-Beyme R (1983) Lung cancer mortality of workers employed in chromate pigment factories. A multicentric European epidemiological study. *J Cancer Res Clin Oncol* **105**: 183-188.

Freudenheim JL, Ritz J, Smith-Warner SA, Albanes D, Bandera EV, van den Brandt PA, Colditz G, Feskanich D, Goldbohm RA, Harnack L, Miller AB, Rimm E, Rohan TE, Sellers TA, Virtamo J, Willett WC, and Hunter DJ (2005) Alcohol consumption and risk of lung cancer: a pooled analysis of cohort studies. *American Journal of Clinical Nutrition* **82**: 657-667.

Fritschi L, Siemiatycki J, and Richardson L (1996) Self-assessed versus expert-assessed occupational exposures. *American Journal of Epidemiology* **144**: 521-527.

Garshick E, Laden F, Hart JE, Rosner B, Smith TJ, Dockery DW, and Speizer FE (2004) Lung cancer in railroad workers exposed to diesel exhaust. *Environmental Health Perspectives* **112**: 1539-1543.

Gelman A, Carlin JB, Stern HS, and Rubin DB. *Bayesian Data Analysis,* NewYork: Chapman and Hall, 1995.

Gibb HJ, Lees PS, Pinsky PF, and Rooney BC (2000) Lung cancer among workers in chromium chemical production. *Am J Ind Med* **38**: 115-126.

Gillis CR, Hole DJ, and Hawthorne VM (1988) Cigarette smoking and male lung cancer in an area of very high incidence. II. Report of a general population cohort study in the West of Scotland. *Journal of Epidemiology & Community Health* **42**: 44-48.

Goldberg MS, Siemiatycki J, and Gerin M (1986) Inter-rater agreement in assessing occupational exposure in a case-control study. *British Journal of Industrial Medicine* **43**: 667-676.

Goodman M, Teta MJ, Hessel PA, Garabrant DH, Craven VA, Scrafford CG, and Kelsh MA (2004) Mesothelioma and lung cancer among motor vehicle mechanics: a meta-analysis. *Annals of Occupational Hygiene* **48**: 309-326.

Goodman SN (1993) P values, hypothesis tests, and likelihood: implications for epidemiology of a neglected historical debate. *American Journal of Epidemiology* **137**: 485-496.

Goodman SN (2001) Of P-values and Bayes: A modest proposal. *Epidemiology* **12**: 295-297.

Gordon T (1974) Hazards in the use of the logistic function. *Journal of Chronic Diseases* **27**: 97-102.

Gotay CC (2005) Behavior and cancer prevention. *Journal of Clinical Oncology* **23**: 301-310.

Greenland S (1987) Interpretation and choice of effect measures in epidemiologic analyses. *American Journal of Epidemiology* **125**: 761-768.

Greenland S (1989a) Comment: cautions in the use of preliminary test estimators. *Statistics in Medicine* **8**: 669-673.

Greenland S (1989b) Modeling and variable selection in epidemiologic analysis. *American Journal of Public Health* **79**: 340-9.

Greenland S (1992) A semi-Bayes approach to the analysis of correlated multiple associations, with an application to an occupational cancer-mortality study. *Statistics in Medicine* **11**: 219-230.

Greenland S (1993) Methods for epidemiologic analyses of multiple exposures: a review and comparative study of maximum-likelihood, preliminary-testing, and empirical-Bayes regression. *Statistics in Medicine* **12**: 717-736.

Greenland S (1994) Hierarchical regression for epidemiologic analyses of multiple exposures. *Environmental Health Perspectives* **102 Suppl 8**: 33-39.

Greenland S (1997) Second-stage least squares versus penalized quasi-likelihood for fitting hierarchical models in epidemiologic analyses. *Statistics in Medicine* **16**: 515-526.

Greenland S (1998) Probability logic and probabilistic induction. *Epidemiology* **9**: 322-332.

Greenland S (1999) Multilevel modeling and model averaging. *Scandinavian Journal of Work, Environment & Health* **25 Suppl 4**: 43-48.

Greenland S (2000a) Principles of multilevel modelling. *International Journal of Epidemiology* **29**: 158-167.

Greenland S (2000b) When should epidemiologic regressions use random coefficients? *Biometrics* **56**: 915-921.

Greenland S (2006) Bayesian perspectives for epidemiological research: I. foundations and basic methods. *Int J Epidemiol* (Advanced electronic publication).

Greenland S and Poole C (1994) Empirical-Bayes and semi-Bayes approaches to occupational and environmental hazard surveillance. *Archives of Environmental Health* **49**: 9-16.

Greenland S and Robins JM (1986) Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology* **15**: 413-419.

Greenland S and Robins JM (1991) Empirical-Bayes adjustments for multiple comparisons are sometimes useful. *Epidemiology* **2**: 244-251.

Greenland S (2000c) Small-sample bias and corrections for conditional maximum-likelihood odds-ratio estimators. *Biostatistics* **1**: 113-122.

Grimsrud TK, Berge SR, Haldorsen T, and Andersen A (2002) Exposure to different forms of nickel and risk of lung cancer. *American Journal of Epidemiology* **156**: 1123-1132.

Grimsrud TK, Berge SR, Haldorsen T, and Andersen A (2005) Can lung cancer risk among nickel refinery workers be explained by occupational exposures other than nickel? *Epidemiology* **16**: 146-154.

Groth DH (1980) Carcinogenicity of beryllium: review of the literature. *Environ Res* **21**: 56-62.

Gustavsson P, Plato N, Lidstrom EB, and Hogstedt C (1990) Lung cancer and exposure to diesel exhaust among bus garage workers. *Scandinavian Journal of Work, Environment & Health* **16**: 348-354.

Hammond EC (1966) Smoking in relation to the death rates of one million men and women. *National Cancer Institute Monograph* **19**: 127-204.

Hanai A, Benn T, Fujimoto I, and Muir CS (1988) Comparison of lung cancer incidence rates by histological type in high and low incidence countries, with reference to the limited role of smoking. *Japanese Journal of Cancer Research* **79**: 445-452.

Harrell FE, Jr., Lee KL, Matchar DB, and Reichert TA (1985) Regression models for prognostic prediction: advantages, problems, and suggested solutions. *Cancer Treat Rep* **69**: 1071-1077.

Harris CC, Vahakangas K, Newman MJ, Trivers GE, Shamsuddin A, Sinopoli N, Mann DL, and Wright WE (1985) Detection of benzo[a]pyrene diol epoxide-DNA adducts in peripheral blood lymphocytes and antibodies to the adducts in serum from coke oven workers. *Proc Natl Acad Sci U S A* **82**: 6672-6676.

Hayes RB (1997) The carcinogenicity of metals in humans. *Cancer Causes Control* **8**: 371-385.

Hayes RB, Lilienfeld AM, and Snell LM (1979) Mortality in chromium chemical production workers: a prospective study. *International Journal of Epidemiology* **8**: 365-374.

Hayes RB, Sheffet A, and Spirtas R (1989) Cancer mortality among a cohort of chromium pigment workers. *American Journal of Industrial Medicine* **16**: 127-133.

Heinrich U (1992) Pulmonary carcinogenicity of cadmium by inhalation in animals. *IARC Scientific Publications (Lyons)* 405-413.

Henderson DW, Rodelsperger K, Woitowitz HJ, and Leigh J (2004) After Helsinki: a multidisciplinary review of the relationship between asbestos exposure and lung cancer, with emphasis on studies published during 1997-2004. *Pathology* **36**: 517-550.

Hernan MA, Hernandez-Diaz S, Werler MM, and Mitchell AA (2002) Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. *American Journal of Epidemiology* **155**: 176-184.

Hernberg S (1996) Significance testing of potential confounders and other properties of study groups--misuse of statistics. *Scandinavian Journal of Work, Environment & Health* **22**: 315-316.

Hertz-Picciotto I and Hu SW (1994) Contribution of cadmium in cigarettes to lung cancer: an evaluation of risk assessment methodologies. *Arch Environ Health* **49**: 297-302.

Hessel PA, Gamble JF, and McDonald JC (2005) Asbestos, asbestosis, and lung cancer: a critical assessment of the epidemiological evidence. *Thorax* **60**: 433-436.

Hill AB and Faning EL (1948) Studies in the incidence of cancer in a factory handling inorganic compounds or arsenic. *British Journal of Industrial Medicine* **5**: 1-6.

Hinds MW, Kolonel LN, and Lee J (1985) Application of a job-exposure matrix to a case-control study of lung cancer. *J Natl Cancer Inst* **75**: 193-197.

Hoar SK, Morrison AS, Cole P, and Silverman DT (1980) An occupation and exposure linkage system for the study of occupational carcinogenesis. *Journal of Occupational Medicine* **22**: 722-726.

Hogstedt C, Andersson K, Frenning B, and Gustavsson A (1982) A cohort study on mortality among long-time employed Swedish chimney sweeps. *Scand J Work Environ Health* **8 Suppl 1**: 72-78.

Holford TR. *Multivariate Methods in Epidemiology,* NewYork: Oxford University Press, 2002.

Holford TR, Zheng T, Mayne ST, Zahm SH, Tessari JD, and Boyle P (2000) Joint effects of nine polychlorinated biphenyl (PCB) congeners on breast cancer risk. *International Journal of Epidemiology* **29**: 975-982.

Horowitz I and Enterline PE (1970) Lung cancer among Jews. *American Journal of Public Health* **60**: 275-282.

Hrubec Z, Blair AE, Rogot E, and Vaught J (1992) Mortality risk by occupation among U.S. veterans of known smoking status, 1954-1980, NIH publication no. 92-3407.US Department of Health and Human Services.

Hsu JC. *Multiple Comparisons: Theory and Methods,* London, UK: Chapman and Hall, 1996.

Hughes JM and Weill H (1991) Asbestosis as a precursor of asbestos related lung cancer: results of a prospective mortality study. *British Journal of Industrial Medicine* **48**: 229-233.

Hughes JM, Weill H, Rando RJ, Shi R, McDonald AD, and McDonald JC (2001) Cohort mortality study of North American industrial sand workers. II. Case-referent analysis of lung cancer and silicosis deaths. *Ann Occup Hyg* **45**: 201-207.

Hung RJ, Brennan P, Malaveille C, Porru S, Donato F, Boffetta P, and Witte JS (2004) Using hierarchical modeling in genetic association studies with multiple markers: application to a case-control study of bladder cancer. *Cancer Epidemiology Biomarkers and Prevention* **13**: 1013-1021.

Hunter D. *The Diseases of Occupations, 5th Edition,* The English Universities Press, Ltd., 1975.

IARC (1977) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Asbestos, Monograph 14." Lyon, France.

IARC (1979) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Some Monomers, Plastics, and Synthetic Elastomers, and Acrolein, Monograph 19." Lyon, France.

IARC (1980) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Some Metals and Metallic Compounds, Monograph 23." Lyon, France.

IARC (1983) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Polynuclear Aromatic Compounds, Part 1, Chemical, Environmental and Experimental Data." Lyon, France.

IARC (1987a) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Overall evaluations of carcinogenicity: an updating of IARC Monographs volumes 1 to 42, Supplement 7." Lyon, France.

IARC (1987b) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Silica and Some Silicates, Monograph 42." Lyon, France.

IARC (1988) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Man-Made Mineral Fibres and Radon, Monograph 43." Lyon, France.

IARC (1989) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Diesel and Gasoline Engine Exhausts and Some Nitroarenes, Monograph 46." Lyon, France.

IARC (1990) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Chromium, Nickel, and Welding, Monograph 49." Lyon, France.

IARC (1992) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Occupational exposures to mists and vapours from strong inorganic acids; and other industrial chemicals, Monograph 54." Lyon, France.

IARC (1993) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Beryllium, Cadmium, Mercury, and Exposures in the Glass Manufacturing Industry, Monograph 58." Lyon, France.

IARC (1995) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Wood dust and formaldehyde, Monograph 62." Lyon, France.

IARC (1997) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Silica, Some Silicates, Coal Dust and Para-Aramid Fibrils, Monograph 68." Lyon, France.

IARC (2001) Ionizing radiation, Part 2: Some internally deposited radionuclides, Monograph 78. Lyon, France: IARC.

IARC (2002) Man-made Vitreous Vibres, Mongraph 81. Lyon, France: IARC.

IARC. Preamble. In: *Tobacco Smoke and Involuntary Smoking, Monograph 83.* Lyon, France: IARC, 2004a.

IARC (2004b) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* "Tobacco Smoke and Involuntary Smoking, Monograph 83." Lyon, France.

ICNCM (1990) Report of the international committee on nickel carcinogenesis in man. *Scandinavian Journal of Work, Environment & Health* **16**: 1-84.

Infante PF, Wagoner JK, and Sprince NL (1980) Mortality patterns from lung cancer and nonneoplastic respiratory disease among white males in the beryllium case registry. *Environ Res* **21**: 35-43.

Infante-Rivard C, Armstrong B, Petitclerc M, Cloutier LG, and Theriault G (1989) Lung cancer mortality and silicosis in Quebec, 1938-85. *Lancet* **2**: 1504-1507.

Ives JC, Buffler PA, and Greenberg SD (1983) Environmental associations and histopathologic patterns of carcinoma of the lung: the challenge and dilemma in epidemiologic studies. *The American Review of Respiratory Disease* **128**: 195-209.

Jarup L (2003) Hazards of heavy metal contamination. *Br Med Bull* **68**: 167-182.

Jarup L and Pershagen G (1991) Arsenic exposure, smoking, and lung cancer in smelter workers--a case-control study. *American Journal of Epidemiology* **134**: 545-551.

Jarup L, Pershagen G, and Wall S (1989) Cumulative arsenic exposure and lung cancer in smelter workers: a dose-response study. *American Journal of Industrial Medicine* **15**: 31-41.

Jostes RF (1996) Genetic, cytogenetic, and carcinogenic effects of radon: a review. *Mutation Research* **340**: 125-139.

Kabat GC (2003) Fifty years' experience of reduced-tar cigarettes: what do we know about their health effects? *Inhalation Toxicology* **15**: 1059-1102.

Kass RE and Steffey D (1989) Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *Journal of the American Statistical Association* **87**: 717-726.

Kauppinen TP, Toikkanen J, and Pukkala E (1998) From cross-tabulations to multipurpose exposure information systems: a new job-exposure matrix. *American Journal of Industrial Medicine* **33**: 409-417.

Kazantzis G, Lam TH, and Sullivan KR (1988) Mortality of cadmium-exposed workers. A five-year update. *Scand J Work Environ Health* **14**: 220-223.

Kazerouni N, Thomas TL, Petralia SA, and Hayes RB (2000) Mortality among workers exposed to cutting oil mist: update of previous reports. *Am J Ind Med* **38**: 410-416.

Kleinbaum DG, Kupper LL, and Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods,* Belmont: Lifetime Learning Publications, 1982.

Kleinbaum DG, Kupper LL, and Muller KE. *Applied Regression Analysis and Other Multivariable Methods, Third Edition,* Pacific Grove, CA: Duxbury Press, 1998.

Korallus U, Ulm K, and Steinmann-Steiner-Haldenstaett W (1993) Bronchical carcinoma mortality in the German chromate-producing industry: the effect of process modification. *Int Arch Occup Environ Health* **65**: 171-178.

Koskela RS, Klockars M, Jarvinen E, Rossi A, and Kolari PJ (1990) Cancer mortality of granite workers 1940-1985. *IARC Scientific Publications (Lyons)* 43-53.

Kristensen P (1992) Bias from nondifferential but dependent misclassification of exposure and outcome. *Epidemiology* **3**: 210-215.

Laden F, Stampfer MJ, and Walker AM (2004) Lung cancer and mesothelioma among male automobile mechanics: a review. *Rev Environ Health* **19**: 39-61.

LaDou J, Landrigan P, Bailar JC, III, Foa V, and Frank A (2001) A call for an international ban on asbestos. *Canadian Medical Association Journal* **164**: 489-490.

Lang JM, Rothman KJ, and Cann CI (1998) That confounded P-value. *Epidemiology* **9**: 7-8.

Langard S, Andersen A, and Gylseth B (1980) Incidence of cancer among ferrochromium and ferrosilicon workers. *British Journal of Industrial Medicine* **37**: 114-120.

Le Cessie S and Houwelingen JC (1992) Ridge estimators in logistic regression. *Applied Statistics* **41**: 191-201.

Lee-Feldstein A (1986) Cumulative exposure to arsenic and its relationship to respiratory cancer among copper smelter employees. *Journal of Occupational Medicine* **28**: 296-302.

Leffondré K, Abrahamowicz M, Siemiatycki J, and Rachet B (2002) Modeling smoking history: a comparison of different approaches. *American Journal of Epidemiology* **156**: 813-823.

Levin ML, Goldstein H, and Gerhardt PR (1950) Cancer and tobacco smoking; a preliminary report. *Journal of the American Medical Association* **143**: 336-338.

Li L, Sun TD, Zhang X, Lai RN, Li XY, Fan XJ, and Morinaga K (2004) Cohort studies on cancer mortality among workers exposed only to chrysotile asbestos: a meta-analysis. *Biomedical and Environmental Sciences* **17**: 459-468.

Liddell FD and Armstrong BG (2002) The combination of effects on lung cancer of cigarette smoking and exposure in quebec chrysotile miners and millers. *Annals of Occupational Hygiene* **46**: 5-13.

Liddell FD and McDonald JC (1980) Radiological findings as predictors of mortality in Quebec asbestos workers. *British Journal of Industrial Medicine* **37**: 257-267.

Lindley DV (1983) Parametric empirical Bayes: theory and application: Comment. *Journal of the American Statistical Association* **78**: 61-62.

Lindley DV (2000) The philosphy of statistics. *The Statistician* **49**: 293-337.

Louis TA (1991) Using empirical Bayes methods in biopharmaceutical research. *Statistics in Medicine* **10**: 811-827.

Lubin JH and Blot WJ (1984) Assessment of lung cancer risk factors by histologic category. *Journal of the National Cancer Institute* **73**: 383-389.

Lubin JH and Boice JD, Jr. (1997) Lung cancer risk from residential radon: meta-analysis of eight epidemiologic studies. *Journal of the National Cancer Institute* **89**: 49-57.

Lubin JH, Pottern LM, Stone BJ, and Fraumeni JF, Jr. (2000) Respiratory cancer in a cohort of copper smelter workers: results from more than 50 years of follow-up. *American Journal of Epidemiology* **151**: 554-565.

Luippold RS, Mundt KA, Dell LD, and Birk T (2005) Low-level hexavalent chromium exposure and rate of mortality among US chromate production employees. *Journal of Occupational and Environmental Medicine* **47**: 381-385.

Machle W and Gregorius F (1948) Cancer of the respiratory system in the United States chromate-producing industry. *Public Health Reports* **63**: 1114-1127.

Magee PN. Organ specificity of chemical carcinogens. In: Marginson GP, ed., *Advances in Medical Oncology, Research and Education. Vol 1.* Proceedings of the 12th International Cancer Congress, Buenos Aires: Pergamon Press, Oxford, 1978.

Maldonado G and Greenland S (1993) Simulation studies of confounder-selection strategies. *American Journal of Epidemiology* **138**: 923-936.

Mancuso TF (1980) Mortality study of beryllium industry workers' occupational lung cancer. *Environ Res* **21**: 48-55.

Mancuso TF (1997a) Chromium as an industrial carcinogen: Part I. *Am J Ind Med* **31**: 129-139.

Mancuso TF (1997b) Chromium as an industrial carcinogen: Part II. Chromium in human tissues. *Am J Ind Med* **31**: 140-147.

Mao Y, Hu J, Ugnat AM, Semenciw R, and Fincham S (2001) Socioeconomic status and lung cancer risk in Canada. *International Journal of Epidemiology* **30**: 809-817.

Markowitz SB and Levin K (2004) Continued epidemic of bladder cancer in workers exposed to ortho-toluidine in a chemical factory. *Journal of Occupational and Environmental Medicine* **46**: 154-160.

Marsh GM, Buchanich JM, and Youk AO (2001a) Historical cohort study of US man-made vitreous fiber production workers: VI. Respiratory system cancer

standardized mortality ratios adjusted for the confounding effect of cigarette smoking. *Journal of Occupational and Environmental Medicine* **43**: 803-808.

Marsh GM, Youk AO, Stone RA, Buchanich JM, Gula MJ, Smith TJ, and Quinn MM (2001b) Historical cohort study of US man-made vitreous fiber production workers: I. 1992 fiberglass cohort follow-up: initial findings. *Journal of Occupational and Environmental Medicine* **43**: 741-756.

Marshall RJ (1991) Mapping disease and mortality rates using empirical Bayes estimators. *J R Stat Soc Ser C Appl Stat* **40**: 283-294.

Mason TJ. Chapter 3: The Descriptive Epidemiology of Lung Cancer. In: Samet JM, ed., *Epidemiology of Lung Cancer*. New York: Marcel Dekker, Inc., 1994.

Masri FA, Comhair SA, Koeck T, Xu W, Janocha A, Ghosh S, Dweik RA, Golish J, Kinter M, Stuehr DJ, Erzurum SC, and Aulak KS (2005) Abnormalities in nitric oxide and its derivatives in lung cancer. *Am J Respir Crit Care Med* **172**: 597-605.

McDonald JC and McDonald AD. Epidemiology of asbestos-related lung cancer. In: Antman K and Aisner J, eds., *Asbestos-related Malignancy*. Orlando: Grune & Straton, 1987.

McDuffie HH, Klaassen DJ, and Dosman JA (1991) Men, women and primary lung cancer--a Saskatchewan personal interview study. *Journal of Clinical Epidemiology* **44**: 537-544.

McLaughlin JK, Hrubec Z, Blot WJ, and Fraumeni JF, Jr. (1995) Smoking and cancer mortality among US veterans: a 26-year followup. *International Journal of Cancer* **60**: 193

Merletti F, Hesseltine E, Saracci R, Simonato L, Vainio HU, and Wilbourn J (1984) Target organs for carcinogenicity of chemicals and undustrial exposures in humans: a review of results in the IARC monographs on the evaluation of carcinogenic risk of chemicals to humans. *Cancer Research* **44**: 2244-2250.

Merlo F, Costantini M, Reggiardo G, Ceppi M, and Puntoni R (1991) Lung cancer risk among refractory brick workers exposed to crystalline silica: a retrospective cohort study. *Epidemiology* **2**: 299-305.

Mickey RM and Greenland S (1989) The impact of confounder selection criteria on effect estimation. *American Journal of Epidemiology* **129**: 125-137.

Miettinen OS (1974) Proportion of disease caused or prevented by a given exposure, trait or intervention. *American Journal of Epidemiology* **99**: 325-332.

Miettinen OS (1976a) Estimability and estimation in case-referent studies. *American Journal of Epidemiology* **103**: 609-620.

Miettinen OS (1976b) Stratification by a multivariate confounder score. *American Journal of Epidemiology* **104**: 609-620.

Miettinen OS. *Theoretical Epidemiology: Principles of Occurrence Research in Medicine,* New York: Wiley, 1985.

Miettinen OS (1999) Etiologic research: Needed revisions of concepts and principles. *Scandinavian Journal of Work, Environment & Health* **25**: 484-490.

Miller AB (1983) Trends in cancer mortality and epidemiology. *Cancer* **51**: 2413-2418.

Mills CA and Porter MM (1950) Tobacco smoking habits and cancer of the mouth and respiratory system. *Cancer Research* **10**: 539-542.

Monson RR and Fine LJ (1978) Cancer mortality and morbidity among rubber workers. *J Natl Cancer Inst* **61**: 1047-1053.

Moore DS (1997) Bayes for beginners? Some reasons to hesitate. *The American Statistician* **51**: 254-261.

Morris CN (1983a) Parametric empirical Bayes: theory and application. *Journal of the American Statistical Association* **78**: 47-65.

Morris CN (1983b) Parametric empirical Bayes: theory and application: Rejoinder. *Journal of the American Statistical Association* **78**: 63-65.

Moulin JJ, Wild P, Haguenoer JM, Faucon D, De Gaudemaris R, Mur JM, Mereau M, Gary Y, Toamain JP, and Birembaut Y (1993) A mortality study among mild steel and stainless steel welders. *British Journal of Industrial Medicine* **50**: 234-243.

Mulvihill JJ (1976) Host factors in human lung tumors: An example of ecogenetics in oncology. *Journal of the National Cancer Institute* **57**: 3-7.

Nafstad P, Haheim LL, Oftedal B, Gram F, Holme I, Hjermann I, and Leren P (2003) Lung cancer and air pollution: a 27 year follow up of 16 209 Norwegian men. *Thorax* **58**: 1071-1076.

Neuberger M and Kundi M (1990) Individual asbestos exposure: smoking and mortality--a cohort study in the asbestos cement industry. *British Journal of Industrial Medicine* **47**: 615-620.

Newhouse ML, Berry G, and Wagner JC (1985) Mortality of factory workers in east London 1933-80. *British Journal of Industrial Medicine* **42**: 4-11.

Newhouse ML and Sullivan KR (1989) A mortality study of workers manufacturing friction materials: 1941-86. *British Journal of Industrial Medicine* **46**: 176-179.

Ng TP (1994) Silica and lung cancer: a continuing controversy. *Ann Acad Med Singapore* **23**: 752-755.

Ng TP, Chan SL, and Lee J (1990) Mortality of a cohort of men in a silicosis register: further evidence of an association with lung cancer. *American Journal of Industrial Medicine* **17**: 163-171.

Nurminen M (1997) On the epidemiologic notion of confounding and confounder identification. *Scandinavian Journal of Work, Environment & Health* **23**: 64-68.

O'Berg MT (1980) Epidemiologic study of workers exposed to acrylonitrile. *Journal of Occupational Medicine* **22**: 245-252.

Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, and Rothschild H (1986) Increased familial risk for lung cancer. *J Natl Cancer Inst* **76**: 217-222.

Ott MG, Holder BB, and Gordon HL (1974) Respiratory cancer and occupational exposure to arsenicals. *Archives of Environmental Health* **29**: 250-255.

Parent ME, Siemiatycki J, and Fritschi L (1998) Occupational exposures and gastric cancer. *Epidemiology* **9**: 48-55.

Parent ME, Siemiatycki J, and Fritschi L (2000) Workplace exposures and oesophageal cancer. *Occupational & Environmental Medicine* **57**: 325-334.

Peduzzi P, Concato J, Kemper E, Holford TR, and Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology* **49**: 1373-1379.

Pershagen G. Chapter Five: Passive Smoking and Lung Cancer. In: Samet JM, ed., *Epidemiology of Lung Cancer*. New York: Marcel-Dekker, Inc., 1994.

Pershagen G, Wall S, Taube A, and Linnman L (1981) On the interaction between occupational arsenic exposure and smoking and its relationship to lung cancer. *Scandinavian Journal of Work, Environment & Health* **7**: 302-309.

Peto R, Lopez AD, Boreham J, Thun M, and Heath CW, Jr. (2005) Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet* **339**: 1268-1278.

PHS (1964) U.S. Public Health Service: Smoking and Health. Report of the advisory committee to the Surgeon General of the Public Health Service. Washington, DC: U.S. Government Printing Office.

Pira E, Pelucchi C, Buffoni L, Palmas A, Turbiglio M, Negri E, Piolatto PG, and La Vecchia C (2005) Cancer mortality in a cohort of asbestos textile workers. *British Journal Of Cancer* **92**: 580-586.

Pisani P, Parkin DM, Bray F, and Ferlay J (1999) Estimates of the worldwide mortality from 25 cancers in 1990. *Int J Cancer* **83**: 18-29.

Poole C (1991) Multiple comparisons? No problem. *Epidemiology* **2**: 241-243.

Pott P (1775) Chirurgical observations relative to the cataract, polypus of the nose, the cancer of the scrotum, the different kinds of ruptures and mortification of the toes and feet. London: Hawes, Clarke, and Collins.

Rachet B, Abrahamowicz M, Sasco AJ, and Siemiatycki J (2003) Estimating the distribution of lag in the effect of short-term exposures and interventions: adaptation of a non-parametric regression spline model. *Statistics in Medicine* **22**: 2335-2363.

Rachet B, Siemiatycki J, Abrahamowicz M, and Leffondre K (2004) A flexible modeling approach to estimating the component effects of smoking behavior on lung cancer. *Journal of Clinical Epidemiology* **57**: 1076-1085.

Ramazzini B. *De Morbis Artificum, translated by Wright, W.C. from the 1713 edition,* New York: Hafner, tr1964.

Richardson S, Hirsch A, Ruffie P, and Bickle M (1987) Relationship between tuberculous scar and carcinomas of the lung. *European Journal of Radiology* **7**: 163-164.

Richiardi L, Forastiere F, Boffetta P, Simonato L, and Merletti F (2005) Effect of different approaches to treatment of smoking as a potential confounder in a case-

control study on occupational exposures. *Occupational & Environmental Medicine* **62**: 101-104.

Robbins H (1955) An empirical Bayes approach to statistics. *Proceedings of the Third Berkeley Symposium on Mathematics, Statistics and Probability*, 1: 157-164.

Robert CP. *The Bayesian Choice, Second Edition,* New York: Springer-Verlag, 2001.

Robins JM and Greenland S (1986) The role of model selection in causal inference from nonexperimental data. *American Journal of Epidemiology* **123**: 392-402.

Rothman KJ (1981) Induction and latent periods. *American Journal of Epidemiology* **114**: 253-259.

Rothman KJ (1990) No adjustments are needed for multiple comparisons. *Epidemiology* **1**: 43-46.

Rothman KJ and Greenland S. *Modern Epidemiology, Second Edition,* Lippincott-Raven, 1998.

Rousseau MC, Straif K, and Siemiatycki J (2005) IARC carcinogen update. *Environmental Health Perspectives* **113**: A580-A581.

Saccomanno G, Archer VE, Auerback R, Saunders P, and Brennan LM (1974) Development of carcinoma of the lung as reflected in exfoliated cells. *Cancer* **34**: 256-270.

Saffioti U. Lung cancer induction by crystalline silica. In: D'Amato R, Slaga T et al, eds., *Relevance of animal studies to the evaluation of human cancer risk.* New York: Wiley-Liss, 1992.

Salvan A, Stayner L, Steenland K, and Smith R (1995) Selecting an exposure lag period. *Epidemiology* **6**: 387-390.

Samet JM. *Epidemiology of Lung Cancer,* New York: Marcel Dekker, Inc., 1994.

Sanderson WT, Ward EM, Steenland K, and Petersen MR (2001) Lung cancer case-control study of beryllium workers. *American Journal of Industrial Medicine* **39**: 133-144.

SAS, Statistical Analysis System Institute, Inc. SAS 8.2. Cary, N.C.: SAS Inc., 1999.

Savitz DA and Olshan AF (1995) Multiple comparisons and related issues in the interpretation of epidemiologic data. *American Journal of Epidemiology* **142**: 904-908.

Schrek R and Baker LA (1950) Tobacco smoking as an etiologic factor in disease; cancer. *Cancer Research* **10**: 49-58.

Sclove SL, Morris C, and Radhakrishna R (1972) Non-optimality of preliminary-test estimators for the mean of a multivariate normal distribution. *Annals of Mathematical Statistics* **43**: 1481-1490.

Seidman H and Selikoff IJ (1990) Decline in death rates among asbestos insulation workers 1967-1986 associated with diminution of work exposure to asbestos. *Annals of the New York Academy of Sciences* **609**: 300-317.

Seidman H, Selikoff IJ, and Gelb SK (1986) Mortality experience of amosite asbestos factory workers: dose-response relationships 5 to 40 years after onset of short-term work exposure. *American Journal of Industrial Medicine* **10**: 479-514.

Selikoff IJ, Hammond EC, and Seidman H (1979) Mortality experience of insulation workers in the United States and Canada, 1943--1976. *Annals of the New York Academy of Sciences* **330**: 91-116.

Shaw GL, Falk RT, Pickle LW, Mason TJ, and Buffler PA (1991) Lung cancer risk associated with cancer in relatives. *Journal of Clinical Epidemiology* **44**: 429-437.

Shields PG (1999) Molecular epidemiology of lung cancer. *Annals of Oncology* **10 supplement 5**: s7-s11.

Shopland DR, Eyre HJ, and Pechacek TF (1991) Smoking-attributable cancer mortality in 1991: Is lung cancer now the leading cause of death among smokers in the United States? *Journal of the National Cancer Institute* **83**: 1142-1148.

Sieber WK, Sundin DS, Frazier TM, and Robinson CF (1991) Development, use, and availability of a job exposure matrix based on national occupational hazard survey data. *American Journal of Industrial Medicine* **20**: 163-174.

Siemiatycki J. Chapter 11: An epidemiological approach to discovering occupational carcinogens by obtaining better information on occupational exposures. In: Harrington MJ, ed., *Recent Advances in Occupational Health.*, 1984.

Siemiatycki J. *Risk Factors for Cancer in the Workplace,* Boston: CRC Press, 1991.

Siemiatycki J (1996) Exposure assessment in community-based studies of occupational cancer. *Occupational Hygiene* **3**: 41-58.

Siemiatycki J (2001) Should Canadian health care professionals support the call for a worldwide ban on asbestos? *Canadian Medical Association Journal* **164**: 495-497.

Siemiatycki J, Day NE, Fabry J, and Cooper JA (1981) Discovering carcinogens in the occupational environment: a novel epidemiologic approach. *Journal of the National Cancer Institute* **66**: 217-225.

Siemiatycki J, DeWar R, Nadon L, and Gerin M (1994) Occupational risk factors for bladder cancer: results from a case-control study in Montreal, Quebec, Canada. *American Journal of Epidemiology* **140**: 1061-1080.

Siemiatycki J, DeWar R, Nadon L, Gerin M, Richardson L, and Wacholder S (1987) Associations between several sites of cancer and twelve petroleum-derived liquids. Results from a case-referent study in Montreal. *Scandinavian Journal of Work, Environment & Health* **13**: 493-504.

Siemiatycki J, DeWar R, and Richardson L (1989) Costs and statistical power associated with five methods of collecting occupation exposure information for population-based case-control studies. *American Journal of Epidemiology* **130**: 1236-1246.

Siemiatycki J, Fritschi L, Nadon L, and Gerin M (1997) Reliability of an expert rating procedure for retrospective assessment of occupational exposures in community-based case-control studies. *American Journal of Industrial Medicine* **31**: 280-286.

Siemiatycki J, Richardson L, Gerin M, Goldberg M, DeWar R, Desy M, Campbell S, and Wacholder S (1986) Associations between several sites of cancer and nine organic dusts: results from an hypothesis-generating case-control study in Montreal, 1979-1983. *American Journal of Epidemiology* **123**: 235-249.

Siemiatycki J, Richardson L, Straif K, Latreille B, Lakhani R, Campbell S, Rousseau M-C, and Boffetta P (2004) Listing occupational carcinogens. *Environmental Health Perspectives* **112**: 1447-1459.

Siemiatycki J, Wacholder S, DeWar R, Cardis E, Greenwood C, and Richardson L (1988) Degree of confounding bias related to smoking, ethnic group, and socioeconomic status in estimates of the associations between occupation and cancer. *Journal of Occupational Medicine* **30**: 617-625.

Simonato L, Vineis P, and Fletcher AC (1988) Estimates of the proportion of lung cancer attributable to occupational exposure. *Carcinogenesis* **9**: 1159-1165.

Sluis-Cremer GK, Liddell FD, Logan WP, and Bezuidenhout BN (1992) The mortality of amphibole miners in South Africa, 1946-80. *British Journal of Industrial Medicine* **49**: 566-575.

Sorahan T, Burges DC, and Waterhouse JA (1987) A mortality study of nickel/chromium platers. *British Journal of Industrial Medicine* **44**: 250-258.

Sorahan T and Esmen NA (2004) Lung cancer mortality in UK nickel-cadmium battery workers, 1947-2000. *Occup Environ Med* **61**: 108-116.

Stanton MF, Layard M, Tegeris A, Miller E, May M, Morgan E, and Smith A (1981) Relation of particle dimension to carcinogenicity in amphibole asbestoses and other fibrous minerals. *J Natl Cancer Inst* **67**: 965-975.

Starr TB, Dalcorso RD, and Levine RJ (1986) Fertility of workers. A comparison of logistic regression and indirect standardization. *American Journal of Epidemiology* **123**: 490-498.

Stayner L, Smith R, Bailer J, Gilbert S, Steenland K, Dement J, Brown D, and Lemen R (1997) Exposure-response analysis of risk of respiratory disease associated with occupational exposure to chrysotile asbestos. *Occupational & Environmental Medicine* **54**: 646-652.

Stayner L, Smith R, Thun M, Schnorr T, and Lemen R (1992) A dose-response analysis and quantitative assessment of lung cancer risk and occupational cadmium exposure. *Annals of Epidemiology* **2**: 177-194.

Steenland K (2005) One agent, many diseases: exposure-response data and comparative risks of different outcomes following silica exposure. *American Journal of Industrial Medicine* **48**: 16-23.

Steenland K and Boffetta P (2000a) Lead and cancer in humans: where are we now? *Am J Ind Med* **38**: 295-299.

Steenland K and Boffetta P (2000b) Lead and cancer in humans: where are we now? *American Journal of Industrial Medicine* **38**: 295-299.

Steenland K, Bray I, Greenland S, and Boffetta P (2000) Empirical bayes adjustments for multiple results in hypothesis-generating or surveillance studies. *Cancer Epidemiology Biomarkers and Prevention* **9**: 895-903.

Steenland K and Brown D (1995) Mortality study of gold miners exposed to silica and nonasbestiform amphibole minerals: an update with 14 more years of follow-up. *American Journal of Industrial Medicine* **27**: 217-229.

Steenland K, Loomis D, Shy CM, and Simonsen N (1996) Review of occupational lung carcinogens. *American Journal of Industrial Medicine* **29**: 474-490.

Steenland K, Mannetje A, Boffetta P, Stayner L, Attfield M, Chen J, Dosemeci M, DeKlerk N, Hnizdo E, Koskela R, and Checkoway H (2001) Pooled exposure-response analyses and risk assessment for lung cancer in 10 cohorts of silica-exposed workers: an IARC multicentre study. *Cancer Causes Control* **12**: 773-784.

Steenland K and Ward E (1991) Lung cancer incidence among patients with beryllium disease: a cohort mortality study. *J Natl Cancer Inst* **83**: 1380-1385.

Steenland NK, Silverman DT, and Hornung RW (1990) Case-control study of lung cancer and truck driving in the Teamsters Union. *American Journal of Public Health* **80**: 670-674.

Stram DO, Huberman M, and Wu AH (2002) Is residual confounding a reasonable explanation for the apparent protective effects of beta-carotene found in epidemiologic studies of lung cancer in smokers? *American Journal of Epidemiology* **155**: 622-628.

Sun GW, Shook TL, and Kay GL (1996) Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* **49**: 907-916.

Takahashi K and Okubo T (1990) A prospective cohort study of chromium plating workers in Japan. *Archives of Environmental Health* **45**: 107-111.

Taylor PR, Qiao YL, Schatzkin A, Yao SX, Lubin J, Mao BL, Rao JY, McAdams M, Xuan XZ, and Li JY (1989) Relation of arsenic exposure to lung cancer among tin miners in Yunnan Province, China. *British Journal of Industrial Medicine* **46**: 881-886.

Teschke K, Olshan AF, Daniels JL, De Roos AJ, Parks CG, Schulz M, Vaughan TL, and Kromhout H (2002) Occupational exposure assessment in case-control studies: opportunities for improvement. *Occupational & Environmental Medicine* **59**: 575-594.

Thomas DC (1985) The problem of multiple inference in identifying point-source environmental hazards. *Environmental Health Perspectives* **62**: 407-414.

Thomas DC and Clayton DG (2004) Betting odds and genetic associations. *J Natl Cancer Inst* **96**: 421-423.

Thomas DC, Goldberg M, DeWar R, and Siemiatycki J (1986) Statistical methods for relating several exposure factors to several diseases in case-heterogeneity studies. *Statistics in Medicine* **5**: 49-60.

Thomas DC, Siemiatycki J, DeWar R, Robins J, Goldberg M, and Armstrong BG (1985) The problem of multiple inference in studies designed to generate hypotheses. *American Journal of Epidemiology* **122**: 1080-1095.

Tukey JW (1991) The philosophy of multiple comparisons. *Statistical Science* **6**: 100-116.

Ugnat AM, Xie L, Semenciw R, Waters C, and Mao Y (2005) Survival patterns for the top four cancers in Canada: the effects of age, region and period. *European Journal of Cancer Prevention* **14**: 91-100.

Viallefont V, Raftery AE, and Richardson S (2001) Variable selection and Bayesian model averaging in case-control studies. *Statistics in Medicine* **20**: 3215-3230.

Vineis P, Airoldi P, Veglia P, Olgiati L, Pastorelli L, Autrup H, Dunning A, Garte S, Gormally E, Hainaut P, Malaveille C, Matullo G, Peluso M, Overvad K, Tjonneland A, Clavel-Chapelon F, Boeing H, Krogh V, Palli D, Panico S, Tumino R, Bueno-de-Mesquita B, Peeters P, Berglund G, Hallmans G, Saracci R, and Riboli E (2005) Environmental tobacco smoke and risk of respiratory cancer and chronic obstructive pulmonary disease in former smokers and never smokers in the EPIC prospective study. *British Medical Journal* **330**: 277-281.

Vineis P, Thomas T, Hayes RB, Blot WJ, Mason TJ, Pickle LW, Correa P, Fontham ET, and Schoenberg J (1988) Proportion of lung cancers in males, due to occupation, in different areas of the USA. *International Journal of Cancer* **42**: 851-856.

Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, and Rothman N (2004) Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* **96**: 434-442.

Wagner JC, Newhouse ML, Corrin B, Rossiter CE, and Griffiths DM (1988) Correlation between fibre content of the lung and disease in east London asbestos factory workers. *British Journal of Industrial Medicine* **45**: 305-308.

Walker C, Everitt J, and Barrett JC (1992) Possible cellular and molecular mechanisms for asbestos carcinogenicity. *American Journal of Industrial Medicine* **21**: 253-273.

Walker SH and Duncan DB (1967) Estimation of the probability of an event as a function of several independent variables. *Biometrics* **54**: 167-179.

Wang JD and Miettinen OS (1982) Occupational mortality studies. Principles of validity. *Scandinavian Journal of Work, Environment & Health* **8**: 153-158.

Ward E, Carpenter A, Markowitz S, Roberts D, and Halperin W (1991) Excess number of bladder cancers in workers exposed to ortho-toluidine and aniline. *J Natl Cancer Inst* **83**: 501-506.

Ward E, Okun A, Ruder A, Fingerhut M, and Steenland K (1992) A mortality study of workers at seven beryllium processing plants. *American Journal of Industrial Medicine* **22**: 885-904.

Weill H (1994) Biological effects: asbestos-cement manufacturing. *Annals of Occupational Hygiene* **38**: 533-8, 413.

Weinberg CR (1993) Toward a clearer definition of confounding. *American Journal of Epidemiology* **137**: 1-8.

Wingo PA, Ries LA, Giovino GA, Miller DS, Rosenberg HM, Shopland DR, Thun MJ, and Edwards BK (1999) Annual report to the nation on the status of cancer, 1973-1996, with a special section on lung cancer and tobacco smoking. *J Natl Cancer Inst* **91**: 675-690.

Wingren G (2004) Mortality and cancer incidence in a Swedish art glassworks--an updated cohort study. *Int Arch Occup Environ Health* **77**: 599-603.

Wingren G and Englander V (1990) Mortality and cancer morbidity in a cohort of Swedish glassworkers. *Int Arch Occup Environ Health* **62**: 253-257.

Winter PD, Gardner MJ, Fletcher AC, and Jones RD (1990) A mortality follow-up study of pottery workers: preliminary findings on lung cancer. *IARC Scientific Publications (Lyons)* 83-94.

Witte JS and Greenland S (1996) Simulation study of hierarchical regression. *Statistics in Medicine* **15**: 1161-1170.

Witte JS, Greenland S, Haile RW, and Bird CL (1994) Hierarchical regression analysis applied to a study of multiple dietary exposures and breast cancer. *Epidemiology* **5**: 612-621.

Witte JS, Greenland S, and Kim LL (1998) Software for hierarchical modeling of epidemiologic data. *Epidemiology* **9**: 563-566.

Wong O (1988) A cohort mortality study of employees exposed to chlorinated chemicals. *American Journal of Industrial Medicine* **14**: 417-431.

World Health Organization (1992) International Programme on Chemical Safety, Environmental Health Criteria 134, "Cadmium".

World Health Organization (1997) International Programme on Chemical Safety, Environmental Health Criteria 188, 2nd edition, "Nitrogen Oxides".

Wu W, Steenland K, Brown D, Wells V, Jones J, Schulte P, and Halperin W (1989) Cohort and case-control analyses of workers exposed to vinyl chloride: an update. *Journal of Occupational Medicine* **31**: 518-523.

Wynder EL and Covey L (1987) Epidemiologic patterns in lung cancer by histologic type. *European Journal of Cancer and Clinical Oncology* **10**: 1491-1496.

Wynder EL and Graham EA (1950) Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma. *Journal of the American Medical Association* **143**: 329-336.

Zhong L, Goldberg MS, Gao YT, and Jin F (1999) Lung cancer and indoor air pollution arising from Chinese-style cooking among nonsmoking women living in Shanghai, China. *Epidemiology* **10**: 488-494.

Appendices

Appendices

1. Letter of ethics approval

2. The occupational chemicals

The following lists include all substances considered in this thesis. Table A-1 lists the chemicals by their identification number, Table A-2 lists them alphabetically, and Table A-3 lists them with descriptive text about the chemicals. The text was abstracted from Siemiatycki (1991), which in turn was based on texts by Parmeggiani (1983) and Brady (1977). Table A-3 also includes the prevalence of each chemical. This was calculated as the proportion of individuals with any level of exposure in their lifetime work history. The text further includes whether or not other chemicals were automatically coded if that particular chemical was coded, the occupations which had the largest numbers of men exposed to that particular chemical, and the top three other correlated chemicals.

Brady GS and Clauser HR. *Materials Handbook, 11th Edition,* New York: McGraw-Hill, 1977.

Parmeggiani L. *Encyclopedia of Occupational Health and Safety, 3rd Edition,* Geneva: International Labor Office, 1983.

Siemiatycki J. *Risk Factors for Cancer in the Workplace,* Boston: CRC Press, 1991.

Table A-1: List of occupational chemicals, by identification number

1. Abrasives Dust
2. Inorganic Insulation Dust
3. Excavation Dust
4. Metallic Dust
5. Asbestos
6. Crystalline Silica
7. Portland Cement
8. Glass Dust
9. Glass Fibres
10. Industrial Talc
11. Brick Dust
12. Clay Dust
13. Concrete Dust
14. Bronze Dust
15. Brass Dust
16. Stainless Steel Dust
17. Mild Steel Dust
18. Inorganic Pigments
19. Mineral Wool Fibres
20. Extenders
21. Aluminium Alloy Dust
22. Ashes
23. Cosmetic Talc
24. Borates
25. Sodium Carbonate
26. Alumina
27. Silicon Carbide
28. Sulfur
29. Calcium Oxide
30. Calcium Sulphate
31. Calcium Carbonate
32. Titanium Dioxide
33. Iron Dust
34. Iron Oxides
35. Copper Dust
36. Zinc Dust
37. Zinc Oxide
38. Lead Oxides
39. Basic Lead Carbonate
40. Lead Chromate
41. Organic Dyes and Pigments
42. Cotton Dust
43. Wool Fibres
44. Wood Dust
45. Grain Dust
46. Flour Dust
47. Fur Dust
48. Hair Dust
49. Starch Dust
50. Sugar Dust

51. Leather Dust
52. Tobacco Dust
53. Natural Rubber
54. Synthetic Fibres
55. Plastic Dust
56. Rayon Fibres
57. Acrylic Fibres
58. Polyester Fibres
59. Nylon Fibres
60. Acetate Fibres
61. Cellulose Nitrate
62. Polyvinyl Chloride
63. Polyvinyl Acetate
64. Poly-Acrylates
65. Alkyds
66. Epoxies
67. Phenol-Formaldehyde
68. Urea-Formaldehyde
69. Polyurethanes
70. Styrene-Butadiene Rubber
71. Polychloroprene
72. Fabric Dust
73. Coal Dust
74. Carbon Black
75. Cellulose
76. Soot
77. Rubber Dust
78. Graphite Dust
79. Hydrogen
80. Carbon Monoxide
81. Hydrogen Cyanide
82. Ammonia
83. Nitrogen Oxides
84. Ozone
85. Hydrogen Fluoride
86. Sulphur Dioxide
87. Hydrogen Sulphide
88. Chlorine
89. Hydrogen Chloride
90. Natural Gas
91. Methane
92. Propane
93. Formaldehyde
94. Acetylene
95. Phosgene
96. Spray Gases
97. Coal Gas
98. Gas Welding Fumes
99. Arc Welding Fumes
100. Soldering Fumes
101. Metal Oxide Fumes
102. Aluminium Fumes

103. Calcium Oxide Fumes
104. Chromium Fumes
105. Manganese Fumes
106. Iron Fumes
107. Nickel Fumes
108. Copper Fumes
109. Zinc Fumes
110. Silver Fumes
111. Tin Fumes
112. Lead Fumes
113. Other Pyrolysis Fumes
114. Cooking Fumes
115. Gasoline Engine Emissions
116. Coal Combustion Products
117. Diesel Engine Emissions
118. Liquid Fuel Combustion Products
119. Wood Combustion Products
120. Natural Gas Combustion Products
121. Jet Fuel Engine Emissions
122. Propane Engine Emissions
123. Plastics Pyrolysis Products
124. Rubber Pyrolysis Products
125. Propane Combustion Products
126. Inorganic Acid Solutions
127. Alkali, Caustic Solutions
128. Javel Water
129. Plating Solutions
130. Nitric Acid
131. Phosphoric Acid
132. Sulphuric Acid
133. Methanol
134. Ethanol
135. Ethylene Glycol
136. Isopropanol
137. Acetic Acid
138. Carbon Tetrachloride
139. Methylene Chloride
140. 1,1,1.-Trichlorethane
141. Trichloroethylene
142. Perchloroethylene
143. Acetone
144. Benzene
145. Toluene
146. Xylene
147. Styrene
148. Phenol
149. Animal and Vegetable Glues
150. Turpentine
151. Linseed Oil
152. Synthetic Adhesives
153. Solvents
154. Waxes, Polishes

155. Leaded Gasoline
156. Kerosene
157. Diesel Oil
158. Heating Oil
159. Mineral Spirits
160. Lubricating Oils and Greases
161. Cutting Fluids
162. Asphalt
163. Coal Tar and Pitch
164. Creosote
165. Hydraulic Fluid
166. Other Mineral Oils
167. Jet Fuel
168. Aviation Gasoline
169. Mineral Spirits+BTX
170. Cutting Fluids pre 1955
171. Cutting Fluids post 1955
172. Other Paints,Varnishes
173. Wood Varnishes, Stains
174. Inks
175. Metal Coatings
176. Cyanides
177. Fluorides
178. Chromium (VI) Compounds
179. Hypochlorites
180. Nitrates
181. Beryllium Compounds
182. Magnesium Compounds
183. Aluminium Compounds
184. Titanium Compounds
185. Vanadium Compounds
186. Chromium Compounds
187. Manganese Compounds
188. Iron Compounds
189. Cobalt Compounds
190. Nickel Compounds
191. Copper Compounds
192. Zinc Compounds
193. Arsenic Compounds
194. Silver Compounds
195. Cadmium Compounds
196. Tin Compounds
197. Antimony Compounds
198. Tungsten Compounds
199. Gold Compounds
200. Mercury Compounds
201. Lead Compounds
202. Alkanes (C18+)
203. Alkanes (C1-C4)
204. Alkanes (C5-C17)
205. Aliphatic Alcohols
206. Aliphatic Aldehydes

207. Chlorinated Alkanes

208. Unsaturated Aliphatic Hydrocarbons

209. Chlorinated Alkenes

210. Aliphatic Esters

211. Aliphatic Ketones

212. Fluorocarbons

213. Glycol Ethers

214. PAH (Any source)

215. PAH (Other)

216. PAH (From wood)

217. PAH (From petroleum)

218. PAH (From coal)

219. Benzo(a)pyrene

220. Monocyclic Aromatic Hydrocarbons

221. Aromatic Alcohols

222. Aromatic Amines

223. Phthalates

224. Isocyanates

225. Cleaning Agents

226. Pharmaceuticals

227. Laboratory Products

228. Fertilizers

229. Pesticides

230. Biocides

231. Bleaches

Table A-2: List of occupational chemicals, alphabetical

Table A-3: List of occupational chemicals, with descriptions

1. *Abrasives Dust.* Dust generated from abrasives during the manufacturing of abrasives or during abrading, smoothing, or polishing of metals, wood, stones, concrete, jewelry, etc. The abrasive could be of a single composition such as silica, or aggregate material containing alumina or silicon carbide, with binders such as vitrified glass, resins or rubber. Main occupations: metal machinists; carpenters; motor vehicle mechanics. Lifetime prevalence: 23.9%. Top three positively correlated chemicals: Alumina (r=0.7), Aluminium Compounds (r=0.7), and Metallic Dust (r=0.5).

2. *Inorganic Insulation Dust.* Dust arising from the placement or removal of any inorganic heat insulating materials including asbestos (chrysotile or amphibole), mineral wool, glass fibers and vermiculite/perlite. Main occupations: pipefitters and plumbers; stationary engineers; carpenters. Lifetime prevalence: 11.2%. Top three positively correlated chemicals: Mineral Wool Fibres (r=0.7), Glass Fibres (r=0.6), and Asbestos (r=0.5).

3. *Excavation Dust.* Dust generated by digging, blasting, drilling, removing or transporting earth or rock for the purpose of mining or quarrying or for the construction of roads, railroads, tunnels and buildings. Main occupations: construction laborers; excavators; truck drivers. Lifetime prevalence: 10.1%. Top three positively correlated chemicals: Crystalline Silica (r=0.6), Concrete Dust (r=0.4), and Portland Cement (r=0.3).

4. *Metallic Dust.* Any metal dusts generated, regardless of the specific metals involved or whether they are known or unknown. Most metals will have undergone a certain amount of surface oxidation but exposure to specific metal oxides (e.g., lead oxides; iron oxides) was coded only when the main exposure was to the oxide itself and not to the metal dust. Main occupations: metal machinists; motor vehicle mechanics; welders and flame cutters. Lifetime prevalence: 27.7%. Top three positively correlated chemicals: Mild Steel Dust (r=0.7), Iron Compounds (r=0.7), and Metal Oxide Fumes (r=0.6).

5. *Asbestos.* A combination of chrysotile and amphibole fibers, both of which are naturally occuring fibrous hydrated silicates. Chrysotile fibres are curly serpentine fibers made up of tiny individual fibrils which take the shape of a spirally wound tube. Chrysotile asbestos is the type most used for textiles, friction materials (e.g. brake and clutch linings) and floor tiles. Main occupations: motor vehicle mechanics; welders and flame cutters; stationary engineers. The amphibole fibers are straight and needle-like silicate structures generally more brittle than chrysotile asbestos fibers. They are useful because of their resistance to heat, wear and corrosion and are generally mixed with chrysotile in asbestos-cement building products, in fire-resistant insulation boards and in other insulation products. Main occupations: stationary engineers; pipefitters and plumbers; electricians. Lifetime prevalence: 16.9%. Top three positively correlated chemicals: Inorg.Insul.Dust (r=0.5), Soot (r=0.4), and Nitrogen Oxides (r=0.4).

6. *Crystalline Silica.* The crystalline forms of free silica are quartz, cristobalite and tridymite. Many sands, clays and rocks are largely composed of small silica crystals; exposure to silica occurred mainly because of sand used in construction, in sand blasting, in foundry molds, clay, glass and stone processing, pottery and brick making. Exposure also occurred to workers involved in mining, quarrying, and rock and soil drilling. Main occupations: carpenters; construction laborers; cabinet and wood furniture makers. Lifetime prevalence: 23.7%. Top three positively correlated chemicals: Excavation Dust (r=0.6), Concrete Dust (r=0.5), and Portland Cement (r=0.4).

7. *Portland Cement.* A powder that can be made into a paste by the addition of water, used in construction for bonding bricks, concrete blocks and stone and for producing concrete slabs, pipes, etc. It consists of about 75% calcium silicates, 5-10% calcium aluminates, 5% calcium

sulphate, 5-10% calcium-aluminium-iron compounds and 1-4% oxides of sodium, potassium and magnesium. Main occupations: construction laborers; stone masons; carpenters. Lifetime prevalence: 7.3%. Top three positively correlated chemicals: Concrete Dust (r=0.5), Crystalline Silica (r=0.4), and Brick Dust (r=0.4).

8. *Glass Dust.* Glass is an inorganic product of fusion which has cooled to a rigid solid without undergoing crystallization. The properties of glasses are determined by their chemical composition and since this can vary infinitely, there are thousands of different glasses available. Potential exposures have been in the construction industry (cutting and installing glass doors and windows) and during grinding, buffing or polishing of optical lenses, prisms, and reflective optics. Main occupations: opticians; glass cutters; glass installers. Lifetime prevalence: 1.4%. Top three positively correlated chemicals: Plastic Dust (r=0.1), Abrasives Dust (r=0.1), and Aluminium Compounds (r=0.1).

9. *Glass Fibres.* Manufactured from the raw ingredients of glass but with processes designed to create fibrous material. These filament fibers have been used mainly as insulators, as plastic reinforcing materials and in special textiles. Main occupations: carpenters; pipefitters and plumbers; motor vehicle mechanics. Lifetime prevalence: 5.9%. Top three positively correlated chemicals: Mineral Wool Fibres (r=0.6), Inorg.Insul.Dust (r=0.6), and Styrene (r=0.4).

10. *Industrial Talc.* Talcs are hydrated magnesium silicates. Depending on where it is mined, commercially available industrial talc is very often geologically associated with other minerals including carbonates, quartz and varying amounts of asbestos minerals (amphiboles and serpentine) in the form of chains or fibres. Industrial grade talcs are widely used as extenders in paints, plastics, ceramic products and paper coatings and in the rubber industry as extenders and dusting powders. Main occupations: painters; motor vehicle mechanics; farmers. Lifetime prevalence: 4.3%. Top three positively correlated chemicals: Titanium Dioxide (r=0.5), Extenders (r=0.5), and Titanium Compounds (r=0.4).

11. *Brick Dust.* Dust generated by the cutting or breaking of bricks, excluding fireclay bricks used for refractory purposes. Included are bricks made from hard burned clay used for buildings, walls and paving, and bricks used for fancy walls which are made with sand or lime. Main occupations: masons; construction laborers; firefighters. Lifetime prevalence: 3.8%. Top three positively correlated chemicals: Concrete Dust (r=0.4), Portland Cement (r=0.4), and Phosgene (r=0.3).

12. *Clay Dust.* Most clays are composed mainly of silica and alumina; they form a paste with water and can be hardened when heated. Clays have been used for making pottery, tiles, bricks, pipes and refractory materials and as extender pigments in paints. Main occupations: foundry molders and coremakers; painters; foundry laborers. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Linseed Oil (r=0.3), Calcium Oxide Fumes (r=0.3), and Iron Fumes (r=0.2).

13. *Concrete Dust.* Dust generated by the cutting, polishing or breaking of concrete which consists of Portland cement, sand, gravel or crushed rock and water. Cast concrete is placed in forms, in a wet state, at the point of use at the construction site and allowed to harden into the form of beams, floor slabs and walls. Main occupations: construction laborers; carpenters; firefighters. Lifetime prevalence: 9.5%. Top three positively correlated chemicals: Crystalline Silica (r=0.5), Portland Cement (r=0.5), and Excavation Dust (r=0.4).

14. *Bronze Dust.* Dust generated when objects made of bronze are cut, abraded, machined, polished, etc. The term bronze is generally applied to any copper alloy where tin is the other major alloying element, although small amounts of other elements are added to modify the characteristics of the bronzes. The product obtained by adding tin to copper is more fusible than

copper and thus better suited for casting. Main occupations: metal machinists; tool and dye makers; foundry molders and coremakers.. Automatics: copper compounds; tin compounds. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Tin Compounds (r=0.4), Silicon Carbide (r=0.3), and Brass Dust (r=0.3).

15. *Brass Dust.* Dust generated when objects made of brass are cut, abraded, machined, polished, etc. Brasses are the most widely used alloys of copper. They are fundamentally binary alloys of copper with zinc but often their properties are modified by addition of other elements in small amounts. Brasses are stronger than copper and are used in structural applications. Uses include bullet jackets, imitation gold jewelry, plumbing hardware, pipes, radiator cases and condenser tubing. Main occupations: metal machinists; machine tool operators; metal grinders. Automatics: copper compounds; zinc compounds. Lifetime prevalence: 2.1%. Top three positively correlated chemicals: Zinc Compounds (r=0.4), Copper Compounds (r=0.4), and Silicon Carbide (r=0.4).

16. *Stainless Steel Dust.* Dust generated when objects made of this metal are cut, abraded, machined, polished, etc. Stainless steel is available in many different compositions but the most common one, usually known as 18-8 stainless steel, is 18% chromium and 8% nickel. Many other elements such as titanium, molybdenum, niobium, silicon and others are also added in small quantities to customize the steel for special purposes. Main occupations: metal machinists; welders and flame cutters; rail transport equipment mechanics. Automatics: chromium compounds; iron compounds; nickel compounds. Lifetime prevalence: 4.5%. Top three positively correlated chemicals: Nickel Compounds (r=0.8), Chromium Compounds (r=0.6), and Chromium Fumes (r=0.5).

17. *Mild Steel Dust.* Dust generated when objects made of this metal are cut, abraded, machined, polished, etc. Mild steel is essentially a combination of iron and carbon (less than 2% carbon). All steels contain manganese (usually at least 0.3%) and small amounts of other metals, which provide the strength and hardness that is required by the construction and manufacturing industries. Main occupations: metal machinists; welders and flame cutters; motor vehicle mechanics. Automatics: manganese compounds; iron compounds. Lifetime prevalence: 16.5%. Top three positively correlated chemicals: Iron Compounds (r=0.8), Metallic Dust (r=0.7), and Abrasives Dust (r=0.5).

18. *Inorganic Pigments.* Insoluble white or colored powders of very fine particle size (0.01 - 1.0 microns) which imparts color and/or other properties (e.g. anti-corrosive properties) to other materials either when mixed intimately with them (dispersion or suspension) or when applied over their surfaces in a thin layer. Inorganic pigments can be subdivided into white (which includes titanium dioxide) , and other colors (iron oxides, lead chromate, etc.) both of which can be of natural or synthetic origin. Organic pigments were coded separately. Main occupations: painters; paper product makers; printshop workers. Lifetime prevalence: 9.2%. Top three positively correlated chemicals: Organic Dyes & Pig. (r=0.6), Extenders (r=0.6), and Alkyds (r=0.6).

19. *Mineral Wool Fibres.* Mineral wool is a glassy fibrous silicate material made by melting and fiberizing slags (slag wool) or natural rocks (rock wool). Mineral wool has been used since the 1930's, mainly as a thermal and acoustical insulator. Main occupations: carpenters; pipefitters and plumbers; stationary engineers. Lifetime prevalence: 6.0%. Top three positively correlated chemicals: Inorg.Insul.Dust (r=0.7), Glass Fibres (r=0.6), and Calcium Sulphate (r=0.4).

20. *Extenders.* A variety of substances used to modify the physical, thermal, mechanical or electrical properties of the products (paints and metal coatings, adhesives, rubber and plastics) to which they are added or to reduce the overall cost of such products. The chemical composition and function of an extender depend on the industry in which the extender is used. Highest

exposure concentrations to extenders occurred during the blending of the raw materials. Main occupations: construction painters; motor vehicle refinishers; motor vehicle mechanics. Lifetime prevalence: 5.6%. Top three positively correlated chemicals: Inorg.Pigments (r=0.6), Alkyds (r=0.6), and Titanium Dioxide (r=0.6).

21. *Aluminium Alloy Dust.* Dust generated when objects made of this alloy are cut, abraded, machined, polished, etc. Pure aluminium (Al) possesses many desirable characteristics: light weight, pleasing appearance, good malleability, high electrical and thermal conductivity and excellent resistance to corrosion. However, in order to achieve the hardness and strength required for industrial use, it must be alloyed with other metals such as copper and magnesium. Aluminium alloys have been widely used for transportation equipment because of their high strength/weight ratios. Main occupations: metal machinists; aircraft assembly workers; welders and flame cutters. Automatics: aluminium compounds. Lifetime prevalence: 5.9%. Top three positively correlated chemicals: Aluminium Compounds (r=0.5), Silicon Carbide (r=0.4), and Cutting Fluids (r=0.4).

22. *Ashes.* The non-combustible residue left after the burning of any susbtance. They contain the residues of all non-volatile substances (e.g. oxides, salts, non-metallic elements) that may have been present in fuel. They are found in largest quantities in industrial processes in which fuel is converted into heat in furnaces, kilns, ovens and boilers. Main occupations: stationary engineers; firefighters; boiler room workers. Lifetime prevalence: 2.9%. Top three positively correlated chemicals: Soot (r=0.5), Phosgene (r=0.5), and Vanadium Compounds (r=0.4).

23. *Cosmetic Talc.* Talcs are hydrated magnesium silicates. Only the pure white talcs (impureties can color it gray, green, brown or red) used in cosmetics and toilet preparations are included in this category. Cosmetic talcs are hand sorted, screened, ground very fine and bolted through silk cloth. Consumer talc products marketed before 1973 may have been contaminated to varying degrees by asbestos. Main occupations: barbers and hairdressers; nurse's aides; physicians and surgeons. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Hair Dust (r=0.5), Ethanol (r=0.3), and Spray Gases (r=0.3).

24. *Borates.* The main substances in this category are borax (Na2B4O7.10H2O), used in the manufacture of special kinds of glass, enamels and glazes, as a scouring and cleansing agent, as a flux in welding and in the soap, leather and cosmetics industries and boric acid (H3BO3), a white, crystalline powder used as a preservative and as a weak antiseptic. Main occupations: jewellers; watch repairmen; dental prosthesis makers. Lifetime prevalence: 1.0%. Top three positively correlated chemicals: Gold Compounds (r=0.5), Silver Fumes (r=0.5), and Silver Compounds (r=0.5).

25. *Sodium Carbonate.* Also known as soda ash, this chemical is an odorless, white, hygroscopic powder considered to be one of the most important industrial alkalis. It occurs naturally but may also be manufactured from salt, ammonia and carbon dioxide. It has been used for cleansing, for softening water, for conditioning boiler feed water (lime-soda process), in glass as a flux to prevent fogging, for refining oils, for the treatment of ores, in the wood-pulp industry, and in soap making. Main occupations: stationary engineers; photographers; textile processors. Lifetime prevalence: 1.6%. Top three positively correlated chemicals: Vanadium Compounds (r=0.4), Alkali, Caustic Solutions (r=0.4), and Beryllium Compounds (r=0.3).

26. *Alumina.* Oxide of aluminium, Al2O3 and its various polymorphs and hydrated species. Corundum (both natural and synthetic), the crystalline form of alumina, has been widely used as an abrasive while the trihydrate is used as an extender pigment in paints, plastics, cosmetics, etc. Main occupations: metal machinists; motor vehicle mechanics; carpenters. Automatics: aluminium compounds. Lifetime prevalence: 15.2%. Top three positively correlated chemicals: Aluminium Compounds (r=0.9), Abrasives Dust (r=0.7), and Silicon Carbide (r=0.5).

27. *Silicon Carbide.* A bluish-black, crystalline, artificial mineral characterized by extreme hardness, a high melting point and chemical inertness. It has been used mainly as an abrasive in the form of granules or powder for shaping, cleaning or polishing surfaces. Other applications include refractories and wear-resistant surfaces. Main occupations: metal machinists; pipefitters and plumbers; motor vehicle mechanics. Lifetime prevalence: 5.5%. Top three positively correlated chemicals: Alumina (r=0.5), Aluminium Compounds (r=0.5), and Abrasives Dust (r=0.4).

28. *Sulfur.* A non-metallic element which exists in several allotropic forms, obtained by the distillation of iron pyrites, as a by-product of metal smelting, and from natural gas. Its most important use (90 %) has been in the production of sulphuric acid. Other uses have included match manufacture, vulcanization of rubber, bleaching agent of paper pulp and wool, and as an agricultural insecticide. Main occupations: farmers; railway mechanics; rubber workers. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Rubber Pyrol.Prod. (r=0.3), Industrial Talc (r=0.2), and Rubber Dust (r=0.2).

29. *Calcium Oxide.* Also known as lime, calcium oxide is produced by calcining limestone (calcium carbonate). Large amounts of lime (or hydrated lime, Ca(OH)2, which was also coded here) have been used in pulp and paper making, as a soil treatment in agriculture and as a 'whitewash' to coat stables, dairies and other farm buildings. It has also been used in masonry mortars, plasters, stucco, and unhairing of skins in leather manufacturing. Main occupations: farmers; stationary engineers; masons. Lifetime prevalence: 7.0%. Top three positively correlated chemicals: Grain Dust (r=0.3), Portland Cement (r=0.3), and Fertilizers (r=0.3).

30. *Calcium Sulphate.* Also known as gypsum, calcium sulphate is a widely distributed naturally occurring mineral. It has been used to produce gypsum wallboard which consists of a core of gypsum sandwiched between two layers of paper. Other uses have been as dental plasters for making tooth impressions, orthopedic plasters, pottery plasters, lamp bases, and patching compounds. Main occupations: painters; carpenters; other construction workers. Lifetime prevalence: 9.8%. Top three positively correlated chemicals: Inorg.Insul.Dust (r=0.4), Mineral Wool Fibres (r=0.4), and Concrete Dust (r=0.4).

31. *Calcium Carbonate.* A mineral occurring naturally in a great variety of calcite rocks which are collectively known as limestone. It has been used as a flux in the melting of iron, as a filler in asphalt, putty, crayons, paints, rubber, plastics and linoleum, for writing on blackboards and as a mild abrasive in polishes. Main occupations: painters and plasterers; teachers; stationary engineers. Lifetime prevalence: 6.2%. Top three positively correlated chemicals: Titanium Dioxide (r=0.3), Extenders (r=0.3), and Titanium Compounds (r=0.3).

32. *Titanium Dioxide.* This extremely dense, powerful opaque white inorganic pigment has great hiding power. It is absolutely inert and therefore permanent. The best quality is produced from ilmenite. It has been used as a pigment in paints, paper, plastics, floor coverings, inks, rubber, ceramics, roofing granules, textiles and as a fluxing agent in welding electrodes. Main occupations: construction painters; motor vehicle refinishers; motor vehicle mechanics. Automatics: titanium compounds. Lifetime prevalence: 3.9%. Top three positively correlated chemicals: Titanium Compounds (r=0.9), Alkyds (r=0.6), and Extenders (r=0.6).

33. *Iron Dust.* Iron (Fe) dust is produced when objects made of iron are cut, abraded, machined, polished, etc. Iron is a silvery white metal, capable of taking a fine polish. It oxidizes readily in the presence of air and water. Exposures to both pure (ingot iron, wrought iron) and cast irons were included here. Cast iron contains from 2 to 4% carbon in the form of graphite or as iron carbide. It has been widely used in automobiles parts such as brake drums, gears, camshafts, hydraulic cylinders, etc., in fireplaces and in kitchen utensils. Iron containing alloys were coded separately. Main occupations: metal machinists; pipefitters and plumbers; metal grinders.

Automatics: iron compounds. Lifetime prevalence: 4.1%. Top three positively correlated chemicals: Iron Compounds (r=0.4), Silicon Carbide (r=0.3), and Cutting Fluids pre 1955 (r=0.3).

34. *Iron Oxides*. All oxides of iron (e.g., Fe3O4, Fe2O3, FeO.Fe2O3). The most important uses of these compounds have been in pigments for plastics, leather, bricks, textiles, paper and concrete products. Exposure to iron oxides is widespread; this also includes exposure to rust. Main occupations: construction painters; motor vehicle mechanics; motor vehicle refinishers. Automatics: iron compounds. Lifetime prevalence: 10.4%. Top three positively correlated chemicals: Iron Compounds (r=0.6), Lead Chromate (r=0.5), and Chromium (VI) Comp. (r=0.4).

35. *Copper Dust*. Dust generated when objects made of copper (Cu) are cut, abraded, machined, polished, etc. Copper is a yellowish red metal which is relatively very malleable. The metal gives a brilliant luster when polished and is second only to silver in electrical conductivity. It has been used in the electrical industry, in water piping, kitchenware, electric motors, coils, and dynamos. Main occupations: machinists and machine tool operators; pipefitters and plumbers; metal grinders. Automatics: copper compounds. Lifetime prevalence: 4.6%. Top three positively correlated chemicals: Copper Compounds (r=0.6), Soldering Fumes (r=0.4), and Lead Fumes (r=0.4).

36. *Zinc Dust*. Dust generated when objects made of or plated with zinc (Zn) are cut, abraded, machined, polished, etc. This bluish-white crystalline metal is obtained from a number of sulphides (the most important one is sphalerite) or oxide ores. It has been used extensively as anticorrosion protection, mainly on steel and iron. It can be applied as a metal by hot dip galvanizing or flame spraying, or used as a pigment in paints. Main occupations: pipefitters and plumbers; sheet-metal workers; machinists and machine tool operators. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Zinc Compounds (r=0.5), Lead Fumes (r=0.5), and Tin Fumes (r=0.4).

37. *Zinc Oxide*. A white, water insoluble powder widely used as a pigment and accelerator in paints and rubbers. In paints it resists the action of ultraviolet light and atmospheric sulphur, and prevents growth of mildew and fungus. Other uses include insulating compounds, sunscreen lotions, and paper coatings. Main occupations: painters; paint mixers; dental prosthesis makers. Automatics: zinc compounds. Lifetime prevalence: 3.2%. Top three positively correlated chemicals: Basic Lead Carb. (r=0.6), Lead Chromate (r=0.6), and Zinc Compounds (r=0.5).

38. *Lead Oxides*. All oxides of lead. Red lead (Pb3O4), and litharge (PbO) which is the yellow lead monoxide, have been used extensively as pigments in paints to protect steel substrates against corrosion. Lead oxides have also been used in the manufacture of glass and in fluxing of earthenware. Lead dioxide (PbO2) is the principle active constituent in the positive plate for lead storage batteries. Main occupations: construction painters; motor vehicle refinishers; ship deck workers. Automatics: lead compounds. Lifetime prevalence: 1.8%. Top three positively correlated chemicals: Alkyds (r=0.3), Inorg.Pigments (r=0.3), and Metal Coatings (r=0.3).

39. *Basic Lead Carbonate*. Commonly known as white lead, this compound is a white, amorphous powder made from metallic lead. It is one of the oldest lead pigments for paints; it has also been used in putty and ceramics. Main occupations: construction painters; pipefitters and plumbers; paint mixers. Automatics: lead compounds. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Linseed Oil (r=0.6), Zinc Oxide (r=0.6), and Lead Chromate (r=0.5).

40. *Lead Chromate*. This category includes not only PbCrO4 itself, commonly known as chrome yellow, but all other other addition compounds containing the lead chromate (e.g., PbO.PbCrO4). These are considered to be the most versatile of the inorganic pigments with a good range of

colors. They are relatively inexpensive, and have been used in wood and metal coatings, printing inks, and as coloring agents in rubber and paper. Main occupations: construction painters; motor vehicle refinishers; motor vehicle mechanics. Automatics: chromium (VI) compounds; chromium compounds; lead compounds. Lifetime prevalence: 3.2%. Top three positively correlated chemicals: Chromium (VI) Comp. (r=0.6), Titanium Dioxide (r=0.6), and Zinc Oxide (r=0.6).

41. *Organic Dyes and Pigments*. Dyes are colored substances which impart their color effects to materials by staining, being absorbed or by chemically reacting; they are used for coloring textiles, leather, paper, plastics, petroleum products and food. Pigments are essentially insoluble in the liquid media in which they are dispersed and are mainly used for paints and plastics. Inorganic pigments were coded separately. Main occupations: construction painters; motor vehicle refinishers; printing press operators. Lifetime prevalence: 8.0%. Top three positively correlated chemicals: Aromatic Amines (r=0.7), Inorg.Pigments (r=0.6), and Lead Chromate (r=0.5).

42. *Cotton Dust*. Dust generated during carding, spinning, weaving, cutting, sewing or handling of cotton or cotton-containing textiles. Cotton is a natural fiber obtained from the Gossypium plant; chemically it is about 90% cellulose and 6% moisture, the remainder being impurities. The textile may have been treated with starches, dyes, inks, sizing or other finishing materials which may have been coded separately. Main occupations: tailors and dressmakers; sewing machine operators; laundry workers and dry cleaners. Lifetime prevalence: 8.8%. Top three positively correlated chemicals: Fabric Dust (r=0.7), Wool Fibres (r=0.6), and Synthetic Fibres (r=0.6).

43. *Wool Fibres*. Dust generated during carding, spinning, weaving, knitting, cutting, sewing and handling of wool or wool-containing textiles. Wool fibers are produced from the hair of sheep or of other animals (goats, llamas). These natural fibers are often blended with synthetic fibers (e.g., acrylic fibers) to make up yarn or textiles. They are often treated with starches, dyes, inks, sizing or other finishing materials, some of which were coded separately. Main occupations: tailors and dressmakers; sewing machine operators; laundry workers and dry cleaners. Lifetime prevalence: 6.1%. Top three positively correlated chemicals: Synthetic Fibres (r=0.7), Polyester Fibres (r=0.7), and Fabric Dust (r=0.6).

44. *Wood Dust*. Generally composed of cellulose, hemicellulose, and lignin but may also include chemicals such as pentachlorophenols and chromated copper arsenate used to improve decay resistance of wood. Wood dust is one of the most common and oldest of occupational exposures. Main occupations: carpenters; cabinet and wood furniture makers; other construction workers. Lifetime prevalence: 22.2%. Top three positively correlated chemicals: Crystalline Silica (r=0.4), Urea-Formald. (r=0.4), and Calcium Sulphate (r=0.4).

45. *Grain Dust*. Dust produced when grains such as wheat, barley and rice are harvested, milled, transported or handled in any other way. The highest exposure concentrations were attributed to grain millers and longshoremen while the lowest exposures were given to farm workers. Lifetime prevalence: 7%. Main occupations: general farmers; dockworkers; dairy farmers. Lifetime prevalence: 7.1%. Top three positively correlated chemicals: Fertilizers (r=0.4), Pesticides (r=0.4), and Calcium Oxide (r=0.3).

46. *Flour Dust*. Dust produced when milled cereals such as wheat, corn, rye, oats, barley and millet are packaged, transported, used for cooking and baking or handled in any other way. Main occupations: bakers; chefs and cooks; dockworkers. Lifetime prevalence: 3.6%. Top three positively correlated chemicals: Cooking Fumes (r=0.3), Sugar Dust (r=0.3), and Natural Gas Comb.Prod. (r=0.2).

47. *Fur Dust.* Dust produced when the furs of aquatic species such as beaver, otter, muskrat and seal or of northern land species such as fox, wolf, mink, weasel, squirrel, bear, badger, marten and raccoon are processed, cut and sewn. Main occupations: furriers; hide and pelt processors; sewing machine operators. Lifetime prevalence: 1.7%. Top three positively correlated chemicals: Leather Dust (r=0.2), Acetate Fibres (r=0.1), and Methanol (r=0.1).

48. *Hair Dust.* Hair is found as a covering and protection on bodies of nearly all mammals, including man. In this study, hair dust was coded for exposure to human hair and to the hair of non fur-producing animals such as cattle, horses, pigs and goats. Main occupations: barbers and hairdressers; upholsterers; hide and pelt processors. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Cosmetic Talc (r=0.5), Spray Gases (r=0.3), and Fluorocarbons (r=0.3).

49. *Starch Dust.* Starch is a soft, white, odorless powder produced from grains such as corn, wheat, rice, potatoes and yams. Starches have been widely used in foodstuffs, adhesives, textile and paper sizing, gelling and thickening agents, and fillers. They have also been used in mining as flocculating agents, in the manufacture of explosives and many chemicals, and as carriers for pigments, inks and dyes. Main occupations: printing press operators; textile weavers; textile winders. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Other Mineral Oils (r=0.2), Cotton Dust (r=0.2), and Phosphoric Acid (r=0.2).

50. *Sugar Dust.* Dust of natural sweeteners that are used in the food and beverage industry. The main sugar used is sucrose, which is a disaccharide of the formula C12H22O11, obtained from sugar cane and beets. Main occupations: bakers; dockworkers; food and beverage processors. Lifetime prevalence: 1.2%. Top three positively correlated chemicals: Flour Dust (r=0.3), Starch Dust (r=0.1), and Alkanes (C1-C4) (r=0.1).

51. *Leather Dust.* Dust generated from skins and hides of animals after they have been cured or tanned by the action of oils, or chemically acted upon by tannins. Leather dust consists of the light, fluffy fibers blown from the buffing and sueding wheels in tanneries or in leather product industries. Main occupations: shoemakers; leather cutters; hide and pelt processors. Lifetime prevalence: 3.1%. Top three positively correlated chemicals: Natural Rubber (r=0.3), Synthetic Adhesives (r=0.3), and Polychloroprene (r=0.2).

52. *Tobacco Dust.* Tobacco is produced from the plant of the genus Nicotiana; it has been used for smoking and chewing, tobacco snuff, as an insecticide, and for production of nicotine. Tobacco dust exposure mainly occurred during the manufacture of cigars, cigarettes, pipe or chewing tobacco or snuff, and to some extent to tobacco farmers. Main occupations: tobacco processors; farmers; other tobacco workers. Lifetime prevalence: 0.8%. Top three positively correlated chemicals: Fertilizers (r=0.1), Pesticides (r=0.1), and Creosote (r=0.1).

53. *Natural Rubber.* Obtained by tapping the bark of the rubber tree and coagulating the milky latex. The rubber is then masticated and blended with various other ingredients such as pigments, vulcanization agents, accelerators, antioxidants, and plasticizers. Often blended with SBR rubber to make automobile tires and other rubber goods, the natural latex may also be used alone in dipped goods such as gloves, toys, and balloons and in adhesives. Main occupations: shoemakers; motor vehicle mechanics; service station attendants. Lifetime prevalence: 4.4%. Top three positively correlated chemicals: Rubber Dust (r=0.9), Styrene-Buta.Rubber (r=0.8), and Rubber Pyrol.Prod. (r=0.4).

54. *Synthetic Fibres.* Dust generated during the manufacturing, spinning, weaving, cutting sewing or handling of artificial or truly synthetic fibers or of textiles containing artificial or synthetic fibers. Artificial fibers are those in which the fiber-forming material is of natural origin (eg., viscose rayon which is regenerated cellulose and celluose acetate fibers) and the true

synthetic fibers are those in which the fiber-forming material is derived from petrochemicals or coal chemicals. They are often treated with starches, dyes, inks, sizing or other finishing materials, some of which were coded separately. Main occupations: tailors and dressmakers; textile cutters; pressers. Lifetime prevalence: 6.4%. Top three positively correlated chemicals: Polyester Fibres (r=0.8), Fabric Dust (r=0.7), and Wool Fibres (r=0.7).

55. *Plastic Dust*. Dust produced when a plastic (of any polymer) material is cut, ground or abraded. It was not coded for paints or adhesives even when these substances produced dusts such as in sanding operations. The main constituents are: polymer resins; color pigments (inorganic and organic), filler pigments, anti-UV agents, plasticizers, fungicides, fire-retardants, stabilizers and anti-static agents. Main occupations: pipefitters and plumbers; shoemakers; dentists. Lifetime prevalence: 5.4%. Top three positively correlated chemicals: Polyvinyl Chloride (r=0.4), Plastics Pyrol.Prod. (r=0.4), and Synthetic Adhesives (r=0.3).

56. *Rayon Fibres*. The oldest man-made fiber, rayon is produced mainly by the viscose process from cellulose. This fiber has been used in wearing apparel (especially in linings and undergarments), car and home upholstering, hospital sanitary products and as tire cord. Main occupations: tailors and dressmakers; upholsterers; pressers. Automatics: synthetic fibers. Lifetime prevalence: 2.5%. Top three positively correlated chemicals: Acetate Fibres (r=0.7), Polyester Fibres (r=0.6), and Nylon Fibres (r=0.6).

57. *Acrylic Fibres*. Includes both acrylic and modacrylic fibers. Acrylic (Orlon®) fibers are made up of long chain polymers composed of at least 85% by weight acrylonitrile units; modacrylic fibers are polymers composed of less than 85% but at least 35% by weight of acrylonitrile units. A relatively new fiber (first produced commercially in 1949), it resembles wool and has often been used as a replacement for it in clothing and upholstery fabrics, yarns and carpets. Main occupations: tailors and sewing machine operators; upholsterers; pressers. Automatics: synthetic fibers. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Polyester Fibres (r=0.7), Nylon Fibres (r=0.6), and Synthetic Fibres (r=0.6).

58. *Polyester Fibres*. Synthetic fibers containing at least 80% of a long-chain polymer composed of an ester of a dihydric alcohol (usually ethylene glycol) and terephthalic acid. They should not be confused with polyester resins which are coded separately. First produced in 1941, polyester (Dacron®) fibers, have been widely used in garments, bedding (e.g., permanent press fabrics), carpets, stuffing for pillows, toys, sleeping bags and comforters as well as for thermal insulation of winter outerwear. Main occupations: tailors and dressmakers; sewing machine operators; textile cutters. Automatics: synthetic fibers. Lifetime prevalence: 4.2%. Top three positively correlated chemicals: Synthetic Fibres (r=0.8), Acrylic Fibres (r=0.7), and Wool Fibres (r=0.7).

59. *Nylon Fibres*. The first truly synthetic fiber, nylon is a manufactured fiber in which the fiber forming substance is any long-chain polyamide having recurring amide groups (HN-C=O) as an integral part of the polymer chain. Nylon (polyamide) fibers, should not be confused with polyamide resins which were coded separately. Type 6 and 6/6 nylon dominate the textile fiber field. More than 60% of the total volume of nylon fibers produced is used in home furnishings, mostly carpets and upholstering. It has also been used in clothing, especially water-resistant outerwear, sweaters and hosiery. Main occupations: textile cutters; sewing machine operators; upholsterers. Automatics: synthetic fibers. Lifetime prevalence: 2.8%. Top three positively correlated chemicals: Synthetic Fibres (r=0.6), Rayon Fibres (r=0.6), and Polyester Fibres (r=0.6).

60. *Acetate Fibres*. These synthetic fibers, introduced in the mid 1920's, are made up of cellulose acetate, a chemical derivative of the naturally occurring polymer cellulose. Included are both types of fibers: acetate, made from partially hydrolyzed cellulose and triacetate, fibers that are fully acetylated. These fibers are mostly used in wearing apparel (women's clothing,

undergarments and linings). Main occupations: tailors and dressmakers; textile cutters; sewing machine operators. Automatics: synthetic fibers. Lifetime prevalence: 1.7%. Top three positively correlated chemicals: Rayon Fibres (r=0.7), Polyester Fibres (r=0.6), and Nylon Fibres (r=0.5).

61. *Cellulose Nitrate.* A thermoplastic resin made by treating cellulose (cotton linters) with a mixture of nitric and sulphuric acids. Cellulose will unite with 1 to 6 molecules to make two types of cellulose nitrate which are both coded under this rubric. The lower nitrates are used for plastics and coatings and the higher nitrates, known as nitrocellulose, for explosives. Main occupations: painters; automobile repairmen; wood furniture makers. Lifetime prevalence: 2.5%. Top three positively correlated chemicals: Aliphatic Esters (r=0.6), Phthalates (r=0.5), and Aliphatic Ketones (r=0.5).

62. *Polyvinyl Chloride.* The general formula for pure PVC is (CH2CHCl)n but some copolymers with polyvinyl acetate are also coded here. The principal markets for these thermoplastic resins have been plumbing pipes and fittings, toys, packaging, flooring and coatings. Main occupations: ship and railway car painters; pipefitters and plumbers; plastic product manufacturing workers. Lifetime prevalence: 1.6%. Top three positively correlated chemicals: Plastic Dust (r=0.4), Plastics Pyrol.Prod. (r=0.3), and Aliphatic Ketones (r=0.2).

63. *Polyvinyl Acetate.* These thermoplastic resins, of general formula (CH3CHCOOCH3)n, are obtained by polymerizing vinyl acetate. They have been mainly used in emulsion-type trade sale paints and in wood adhesives. Other end uses include textile and paper coatings. Main occupations: construction painters; carpenters and wood cabinet makers. Lifetime prevalence: 2.8%. Top three positively correlated chemicals: Wood Varnishes, Stains (r=0.4), Poly-Acrylates (r=0.4), and Mercury Compounds (r=0.4).

64. *Poly-Acrylates.* Thermoplastic resins produced by polymerization of the esters of acrylic or methacrylic acid. The most important resin is polymethylmethacrylate (sheet plastic produced with this resin is commonly know as plexiglass). First produced commercially in 1931, acrylic plastics have been widely used as a substitute for glass and for dental prosthesis. Acrylic resins have also been widely used in coatings, most notably in water-based trade sale paints and in motor vehicle paints. Main occupations: construction painters; motor vehicle refinishers; dentists and dental prosthesis makers. Lifetime prevalence: 2.9%. Top three positively correlated chemicals: Titanium Dioxide (r=0.6), Titanium Compounds (r=0.5), and Lead Chromate (r=0.5).

65. *Alkyds.* Thermosetting oil-modified polyester resins made by the esterification of a polybasic acid with a polyhydric alcohol; the resins are reacted with oils, fatty acids or other resins. They have been used extensively in solvent-based coatings (especially trade sale paints) and in printing inks. Main occupations: painters, automobile repairmen; paint mixers. Lifetime prevalence: 4.2%. Top three positively correlated chemicals: Titanium Dioxide (r=0.6), Extenders (r=0.6), and Inorg.Pigments (r=0.6).

66. *Epoxies.* Thermosetting synthetic resins characterized by a highly strained triangular ring consisting of an oxygen atom bonded to two adjoining and bonded carbon atoms. They are usually made by reacting epichlorohydrin with polyhydroxy compounds (very often bisphenol A). Excellent mechanical and electrical properties have made these resins very useful in adhesives, resistant coatings and electrical insulating materials. Main occupations: painters; watch and clock repairmen; electric and electronic equipment operators. Lifetime prevalence: 1.2%. Top three positively correlated chemicals: Borates (r=0.2), Cadmium Compounds (r=0.2), and Methylene Chloride (r=0.2).

67. *Phenol-Formaldehyde.* Thermosetting resins prepared by reacting phenol with formaldehyde; they are among the oldest of the synthetic plastic materials, dating back to 1909. They have been widely used in adhesives (especially outdoor plywood bonding), foundry molds and cores (as a binder for sand), brake linings (as a binder for the asbestos fibers), coatings (electrical insulating varnishes), resin-bonded grinding wheels, laminates, thermal and acoustical insulation materials (as a binding agent for glass fibers or mineral wool) and in castings. Main occupations: carpenters; cabinet and wood furniture makers; electric and electronic equipment fabricators. Lifetime prevalence: 3.8%. Top three positively correlated chemicals: Urea-Formald. (r=0.6), Phenol (r=0.4), and Aromatic Alcohols (r=0.3).

68. *Urea-Formaldehyde.* These resins are among the most widely used of the amino resins (resins produced by the addition reaction between formaldehyde and such compounds as urea, melamine, aniline, ethylene, sulphonamide). They are used extensively as plywood and particleboard adhesives, especially for indoor use (furniture particleboard, indoor use plywood). In the past they were also used for imparting wrinkle recovery to cellulosic fabrics but have been gradually replaced by melamine resins. Urea-formaldehyde resins are still widely used in molded plastics (especially electrical wall plates and connectors, lighting fixtures and reflectors) and in oven-cured industrial finish coatings (porcelain-type finish). Main occupations: carpenters; cabinet and wood furniture makers; adhesives mixers. Lifetime prevalence: 4.3%. Top three positively correlated chemicals: Phenol-Formald. (r=0.6), Synthetic Adhesives (r=0.4), and Wood Dust (r=0.4).

69. *Polyurethanes.* Generally considered to cover all products of reaction between isocyanates and polyhydroxy compounds. These resins have been used in a great variety of products: rigid and flexible foam products, hard and soft plastic products, elastomers, paints, varnishes and adhesives. Main occupations: motor vehicle refinishers; upholsterers; textile workers. Lifetime prevalence: 1.6%. Top three positively correlated chemicals: Isocyanates (r=0.5), Titanium Dioxide (r=0.2), and Poly-Acrylates (r=0.2).

70. *Styrene-Butadiene Rubber.* SBR is a copolymer of 1,3-butadiene and styrene (with butadiene furnishing at least 50% of the polymer units). It is a general purpose rubber often blended with natural rubber. Tire manufacturing represents the largest single rubber application and the largest end use of SBR rubber. Water-based trade sale paints were formulated with SBR in the early 50's. Main occupations: construction painters; motor vehicle mechanics; shoemakers and repairmen. Lifetime prevalence: 4.2%. Top three positively correlated chemicals: Natural Rubber (r=0.8), Rubber Dust (r=0.7), and Zinc Oxide (r=0.3).

71. *Polychloroprene.* The world's entire production of chloroprene (CH2=CClCHCH2) is used in the manufacture of polychloroprene latex and rubber (also called neoprene). These elastomers have been widely used in synthetic adhesives, mainly solvent-based contact cements. Their excellent resistance to oils, chemicals and sunlight have also led to their use in a variety of rubber products such as conveyer belts, footwear, hose covers and wire coverings. Main occupations: carpenters; shoemakers; cabinet and wood furniture makers. Lifetime prevalence: 3.1%. Top three positively correlated chemicals: Synthetic Adhesives (r=0.4), Toluene (r=0.4), and Urea-Formald. (r=0.3).

72. *Fabric Dust.* Dust generated during cutting, sewing or handling of fabrics made of either natural or synthetic fibers. The fabric may contain pigments used to print or dye the fabric, sizing or other finishing materials which may contribute to the toxicity of the dust and which may have been coded separately. Main occupations: tailors and dressmakers; textile cutters; textile shippers and material handlers. Lifetime prevalence: 9.8%. Top three positively correlated chemicals: Cotton Dust (r=0.7), Synthetic Fibres (r=0.7), and Wool Fibres (r=0.6).

73. *Coal Dust.* Coal is composed mainly of carbon with smaller amounts of hydrogen, nitrogen, oxygen, sulphur and other organic aromatic compounds. In the past, it was used primarily as a fuel. The greatest exposure to coal dust occurs among miners and others who handled it in its raw form. Main occupations: stationary engineers; truck drivers (coal delivery); coal miners. Lifetime prevalence: 5.4%. Top three positively correlated chemicals: Coal Comb.Products (r=0.5), PAH (Coal) (r=0.4), and Ashes (r=0.3).

74. *Carbon Black.* An amorphous powdered carbon resulting from the incomplete combustion of liquid or gaseous hydrocarbons in a limited air supply. It contains essentially 88-95% elemental carbon, 0.4-11% oxygen and 0.05-0.8% hydrogen. Includes all types: furnace black, thermal black, channel black, etc. It has been used mainly to reinforce rubber for tires and other rubber articles and as a pigment in inks and paints. Main occupations: rubber molders and mixers; printers and typesetters; painters. Lifetime prevalence: 5.0%. Top three positively correlated chemicals: Inorg.Pigments (r=0.6), Aromatic Amines (r=0.5), and Lead Chromate (r=0.4).

75. *Cellulose.* The main constituent of the cell walls of plants. Industrial cellulose is made from wood or cotton pulp. It is used for paper making but also as a starting material for cellulose acetate and cellulose nitrate. Exposure has been mainly coded to workers exposed to paper fibres. Main occupations: material handlers; paper product manufacturing workers; shippers and receivers. Lifetime prevalence: 5.8%. Top three positively correlated chemicals: Inks (r=0.3), Animal & Vege.Glues (r=0.2), and Inorg.Pigments (r=0.2).

76. *Soot.* A black carbonaceous substance formed by the combustion of coal, wood, oil or other fuel. In addition to carbon and PAHs, it may contain other mineral constituents as well as trace amounts of metals (e.g., lead, vanadium, barium, chromium). The composition of soot varies according to the fuel and the completeness of the combustion. Main occupations: motor vehicle mechanics; stationary engineers; firefighters. Lifetime prevalence: 8.5%. Top three positively correlated chemicals: Benzo(a)pyrene (r=0.6), Ashes (r=0.5), and Sulphur Dioxide (r=0.5).

77. *Rubber Dust.* Includes rubber dusts of both natural and synthetic origin. Whenever possible, the specific type of rubber used was also coded (mainly styrene-butadiene rubber or natural rubber). Highest exposures occurred during buffing of tires in recapping operations; exposure was also coded often to shoe repairmen and shoemakers because of rubber sole buffing. Main occupations: Motor vehicle mechanics; shoemakers and repairmen; dental prosthesis makers. Lifetime prevalence: 3.7%. Top three positively correlated chemicals: Natural Rubber (r=0.9), Styrene-Buta.Rubber (r=0.7), and Rubber Pyrol.Prod. (r=0.4).

78. *Graphite Dust.* Graphite is a form of carbon in which carbon atoms are arranged in a hexagonal, layerlike crystalline structure. Natural graphite, which has a grayish black color, has been used for pencils, as a stove polish, in foundry mold facings, in packing seals and as a lubricant. The artificial variety may be made by heating carbon to a temperature of nearly 3000°C and holding for a sufficient time for formation of the orderly hexagonal crystal pattern. Main occupations: electricians; refractory brick layers and repairmen; aluminium refinery workers. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Phenol (r=0.3), Phenol-Formald. (r=0.2), and Aromatic Alcohols (r=0.2).

79. *Hydrogen.* The lightest known substance. It is one of the main consitituents of coal gas. It is a colorless, odorless gas used as a rocket fuel, a welding fuel, a reducing agent, a reagent in various organic synthesis, a hydrogenating agent for vegetable or animal oils and as a raw material in the manufacturing of ammonia and hydrogen chloride. Occupational exposures have mainly occured due to hydrogen being liberated as an unwanted by-product in industrial processes. Main occupations: electroplaters; chefs and cooks; pipefitters and plumbers. Lifetime prevalence: 2.1%. Top three positively correlated chemicals: Coal Gas (r=0.7), Plating Solutions (r=0.4), and Methane (r=0.4).

80. *Carbon Monoxide.* A colorless, tasteless and almost odorless gas which is lighter than air and burns in air with a blue flame. It is an active reducing agent for chemicals at elevated temperatures, but is mostly encountered as a waste product of incomplete combustion of carbonaceous material. Potential sources of carbon monoxide exposure include engine emissions (gasoline diesel, jet fuel, etc.), foundry furnaces and other industrial furnaces, welding operations, etc. Main occupations: motor vehicle drivers; motor vehicle mechanics and repairmen; welders and flame cutters. Lifetime prevalence: 50.0%. Top three positively correlated chemicals: Gas Eng.Emissions (r=0.7), Lead Compounds (r=0.7), and PAH (Any) (r=0.7).

81. *Hydrogen Cyanide.* A colorless gas with the characteristic smell of bitter almonds. It has been used in extraction of gold, in extermination of rodents and insects in orchards and tobacco farms, in metallurgy and in jewelry manufacturing. Occupational exposure also occurred to electroplaters (cyanide solutions) and firefighters (plastic pyrolysis). Main occupations: firefighters; electroplaters; jewellers. Automatics: cyanides. Lifetime prevalence: 1.7%. Top three positively correlated chemicals: Cyanides (r=0.9), Phosgene (r=0.5), and Isocyanates (r=0.5).

82. *Ammonia.* A by-product of coal distillation and is also produced by passing nitrogen, hydrogen and a catalyst through an electric arc. It is an important source of various nitrogen containing compounds. An enormous quantity of ammonia is used in the production of fertilizers. As a gas it has been used in refrigeration and in nitriding, bright annealing, and for sintering metals. As an aqueous solution (NH4OH), it has been used in the textile and pharmaceutical industries, in medicine, in trade sale paints, in fire extinguishers and in consumer cleaning products. Main occupations: janitors; painters; firefighters. Lifetime prevalence: 10.5%. Top three positively correlated chemicals: Biocides (r=0.4), Waxes, Polishes (r=0.3), and Cleaning Agents (r=0.3).

83. *Nitrogen Oxides.* Formed when nitrogen is oxidized in a high temperature flame, an electric arc or an internal combustion engine. The source of the nitrogen is often the atmosphere itself of which nitrogen is a major constituent. The oxides are rarely released pure into the atmosphere, but occur as mixtures, the composition of which depends upon the source and the local conditions. Main occupations: welders and flame cutters; motor vehicle mechanics; pipefitters and plumbers. Lifetime prevalence: 21.6%. Top three positively correlated chemicals: Sulphur Dioxide (r=0.7), Metal Oxide Fumes (r=0.5), and Iron Fumes (r=0.5).

84. *Ozone.* A bluish gas with a slightly pungent odor. It is generated from oxygen on exposure to ultraviolet radiation and in the vicinity of electrical sources. It is a powerful oxidizer capable of breaking down most organic compounds. The potential sources of exposure to ozone in industry are leakages from ozone-using processes, high voltage electrical equipment, electric arc welding, electric furnaces, photocopying machines, and in the bleaching of textiles, paper pulp, starch and sugar. Main occupations: welders and flame cutters; mechanics; pipefitters and plumbers. Lifetime prevalence: 6.2%. Top three positively correlated chemicals: Calcium Oxide Fumes (r=0.7), Iron Fumes (r=0.6), and Manganese Fumes (r=0.6).

85. *Hydrogen Fluoride.* Anhydrous hydrogen fluoride is a colorless gas prepared by the action of sulphuric acid on calcium fluoride. It is strongly corrosive and irritating. Aqueous solution and salts of hydrofluoric acid are used in the production of fluorides and plastics, in frosting and etching glass, in polishing crystals, in enameling and galvanizing iron, in working silk, in analytic chemistry, and to increase the porosity of ceramics. In our study, many exposures resulted from thermal degradation of fluoride coatings on welding electrodes Main occupations: welders and flame cutters; sheet metal workers; pipefitters and plumbers. Automatics: fluorides. Lifetime prevalence: 2.9%. Top three positively correlated chemicals: Fluorides (r=0.9), Ozone (r=0.5), and Manganese Fumes (r=0.5).

86. *Sulphur Dioxide.* A colorless, non-flammable gas with a pungent odor. It can be manufactured by the combustion of sulphur, the roasting of sulphides or the calcining of sulphates. It is an intermediate in the production of sulphuric acid and is used as a bleaching agent in various industries. However most exposures occur when the gas is released as an unwanted by-product of fuel-burning operations or in the smelting of sulphide ores. Main occupations: foundry workers; stationary engineers; welders. Lifetime prevalence: 15.5%. Top three positively correlated chemicals: Nitrogen Oxides (r=0.7), Benzo(a)pyrene (r=0.5), and Soot (r=0.5).

87. *Hydrogen Sulphide.* A flammable, poisonous gas which has the characteristic smell of rotten eggs. It occurs naturally as a decomposition product of metal sulphides and other organic matter in mines, springs and sewers. It is also a by-product of many chemical processes involving rayon, rubber, petroleum products, leather, and coke production. Main occupations: miners; quarry workers; farmers. Lifetime prevalence: 4.5%. Top three positively correlated chemicals: Phosgene (r=0.3), Sulphur Dioxide (r=0.3), and Rubber Pyrol.Prod. (r=0.3).

88. *Chlorine.* This gas has a distinctive, irritating odor and a yellowish-green color. It is mainly produced commercially by electrolysis of brine. Chlorine has been used for the production of bleaching powders, the treatment of water supplies and refuse and for chlorination in swimming pools. Exposure to chlorine may also occur when hypochlorites are used as bleaches and cleaning agents. Main occupations: launderers; textile processors; pipefitters and plumbers. Lifetime prevalence: 2.4%. Top three positively correlated chemicals: Phosgene (r=0.5), Hydrogen Cyanide (r=0.4), and Cyanides (r=0.4).

89. *Hydrogen Chloride.* This colorless gas, heavier than air, may be used directly as a catalyst. However, the aqueous solution of hydrogen chloride, known as hydrochloric or muriatic acid, is more commonly used industrially for pickling and cleaning metal parts, in the production of glues, in the manufacture of chlorine and pharmaceuticals, for tanning, etching, and for treating oils and fats. Hydrogen chloride is also an unwanted contaminant in certain operations such as plastic pyrolysis (e.g., firefighters) and galvanizing. Main occupations: pipefitters and plumbers; bricklayers; jewellers. Lifetime prevalence: 6.9%. Top three positively correlated chemicals: Inorg.Acid Solutions (r=0.5), Hydrogen Cyanide (r=0.4), and Cyanides (r=0.4).

90. *Natural Gas.* A mixture of light aliphatic hydrocarbon gases, chiefly methane, in increasing demand as a fuel because of its low sulphur content. It has been distributed in the Montreal area since 1957. Mainly used in power plants, industrial process heating and space heaters, the highest exposure in our study occurred among pipe fitters and plumbers repairing gas leaks. It was also frequently coded as a background exposure for those exposed to the combustion products of natural gas (e.g., cooks). Main occupations: chefs and cooks; restaurant managers; stationary engineers. Automatics: alcanes (C1-C4); methane. Lifetime prevalence: 2.6%. Top three positively correlated chemicals: Methane (r=0.7), Natural Gas Comb.Prod. (r=0.7), and Alkanes (C1-C4) (r=0.5).

91. *Methane.* A colorless odorless gas, it is the principal constituent of natural gas. It is sometimes known as marshgas (decomposition of natural organic matter) and as firedamp in coal mines. Methane has been used as a fuel for cooking and heating, and as a raw material for many synthetic products such as formaldehyde, acetylene, and hydrogen cyanide. Main occupations: chefs and cooks; restaurant managers; coal miners. Automatics: alkanes (C1-C4). Lifetime prevalence: 4.5%. Top three positively correlated chemicals: Natural Gas (r=0.7), Alkanes (C1-C4) (r=0.7), and Natural Gas Comb.Prod. (r=0.5).

92. *Propane.* A hydrocarbon used as a raw material for the production of several chemicals such as propylene, hydrogen and perchloroethylene. It is used extensively by pipefitters and plumbers as a welding fuel and as fuel in restaurants and laboratories. Main occupations: chefs and cooks;

pipefitters and plumbers; restaurant managers. Automatics: alkanes (C1-C4). Lifetime prevalence: 3.8%. Top three positively correlated chemicals: Propane Comb.Prod. (r=0.9), Alkanes (C1-C4) (r=0.6), and Lead Fumes (r=0.3).

93. *Formaldehyde.* A colorless gas obtained by the oxidation of methyl alcohol, it is marketed as a 37% solution by weight under the name of formalin. Formaldehyde has been mainly used for plastics and resin manufacture (see urea-formaldehyde, melamine-formaldehyde and phenol formaldehyde), as a disinfectant and fumigant and as a preservative and hardener of tissues in embalming fluids. Exposure to formaldehyde in the workplace can result from the use of formaldehyde gas or formaldehyde solutions, from outgassing or thermal decomposition of formaldehyde resins or from thermal decomposition of other resins, plastics or organic materials. Main occupations: carpenters; textile workers; laundry workers and dry cleaners. Automatics: aliphatic aldehydes. Lifetime prevalence: 14.3%. Top three positively correlated chemicals: Aliphatic Aldehydes (r=0.9), Urea-Formald. (r=0.3), and Phenol-Formald. (r=0.2).

94. *Acetylene.* A colorless gas obtained by reaction of calcium carbide and water or produced from petroleum; it is normally marketed compressed in cylinders. It is used extensively for welding and flame cutting of metals and as a starting material for other chemicals. Main occupations: welders and flame cutters; motor vehicle mechanics; pipefitters and plumbers. Automatics: unsaturated aliphatic hydrocarbons. Lifetime prevalence: 4.3%. Top three positively correlated chemicals: Unsat.Aliph.Hydrocarb. (r=0.9), Gas Welding Fumes (r=0.6), and Iron Fumes (r=0.5).

95. *Phosgene.* A colorless gas at room temperature which is made by the action of chlorine on carbon monoxide. Once used as a war gas, it was later used mainly in the manufacture of many organic chemicals. Exposure mainly occurs as a product of combustion of volatile chlorinated solvents. Main occupations: firefighters, welders and flame cutters; dry cleaners. Lifetime prevalence: 1.5%. Top three positively correlated chemicals: Hydrogen Cyanide (r=0.5), Cyanides (r=0.5), and Chlorine (r=0.5).

96. *Spray Gases.* Gaseous propellants used to form aerosols with liquids or solids. The most common spray gases are chlorofluorocarbons, chlorinated hydrocarbons, propane, butane, vinylchloride (not used as a propellant since 1974) and methylene chloride. Many personal hygiene products (hair sprays, shaving lather, deodorants and antiperspirants) contain spray gases. Main occupations: barbers and hairdressers; motor vehicle mechanics; janitors. Lifetime prevalence: 1.8%. Top three positively correlated chemicals: Fluorocarbons (r=0.5), Hair Dust (r=0.3), and Cosmetic Talc (r=0.3).

97. *Coal Gas.* Produced by heating coal, usually in the presence of steam, at temperatures in excess of 750°C. The composition of the resulting gas can vary but contains roughly 50-53% hydrogen, 25-30% methane and 5-10% carbon monoxide; it has been used as a fuel for heating homes and for cooking in restaurants. Main occupations: chefs and cooks; welders and flame cutters; bakers. Automatics: alkanes (C1-C4); carbon monoxide; hydrogen; methane; hydrogen sulphide. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Hydrogen (r=0.7), Methane (r=0.5), and Alkanes (C1-C4) (r=0.3).

98. *Gas Welding Fumes.* Any fumes generated during the joining or cutting of metals using gas welding techniques. In this process the heat of fusion is obtained from combustion of oxygen and one of several gases such as acetylene, methylacetylene-propadiene (MAPP), propane, and hydrogen. Gas welding fumes include metal fumes from base and filler metals, fumes from the fluxes and from the combustible gases used. Main occupations: welders and flame cutters; motor vehicle mechanics; industrial equipment mechanics. Lifetime prevalence: 10.8%. Top three positively correlated chemicals: Metal Oxide Fumes (r=0.7), Arc Welding Fumes (r=0.6), and Acetylene (r=0.6).

99. *Arc Welding Fumes*. Any fumes generated during the joining or cutting of metals using arc welding techniques. In this process the heat of fusion is obtained by striking an electric arc between an electrode and the metal workpiece. The fumes coded here would include those generated from the base metal, from the electrodes and from electrode coverings (which may contain inorganic and organic fluxing compounds) and/or their decomposition products. Main occupations: welders and flame cutters; industrial equipment mechanics; motor vehicle mechanics. Lifetime prevalence: 10.9%. Top three positively correlated chemicals: Metal Oxide Fumes (r=0.6), Gas Welding Fumes (r=0.6), and Ozone (r=0.6).

100. *Soldering Fumes*. Fumes generated during soldering operations. Soldering is the joining of metal using a filler metal (solder) with a melting point less than 400°C. The quality of the fumes depends on the composition of the solder and of the fluxes and on the production techniques used. Main occupations: pipefitters and plumbers; electric and electronic equipment repairmen; construction electricians. Lifetime prevalence: 6.3%. Top three positively correlated chemicals: Tin Compounds (r=0.7), Lead Fumes (r=0.5), and Tin Fumes (r=0.5).

101. *Metal Oxide Fumes*. Any oxidized metal fumes formed during high temperature treatment of metals in industrial operations such as welding, casting, smelting, etc. Exposure to metal oxide fumes was also coded when metallic compounds (e.g. thermal decomposition of paints containing inorganic pigments during welding operations) are vaporized. Main occupations: welders and flame cutters; pipefitters and plumbers; motor vehicle mechanics. Lifetime prevalence: 19.1%. Top three positively correlated chemicals: Gas Welding Fumes (r=0.7), Arc Welding Fumes (r=0.6), and Iron Fumes (r=0.6).

102. *Aluminium Fumes*. Fumes produced during high temperature processes involving aluminium (Al)-containing alloys or ores. The major use of aluminium has been structural, in the building, aircraft, and automotive industries; it has also been used in housewares, and in containers and packaging. Exposure to aluminium fumes has generally occurred during welding, casting or smelting and refining operations. Main occupations: welders and flame cutters; aluminium refinery workers; motor vehicle repairmen. Automatics: aluminium compounds. Lifetime prevalence: 2.1%. Top three positively correlated chemicals: Copper Fumes (r=0.4), Fluorides (r=0.4), and Chromium Fumes (r=0.4).

103. *Calcium Oxide Fumes*. Fumes generated during high temperature processes involving calcium oxide (CaO). Exposure occurs during steel making where slag forming materials such as calcium oxide and dolomite are added to the charge as fluxing agents and in certain welding operations because carbon steel electrode coverings may contain calcium oxide. Main occupations: welders and flamecutters; mechanics; foundry molders and coremakers. Lifetime prevalence: 6.5%. Top three positively correlated chemicals: Iron Fumes (r=0.8), Manganese Fumes (r=0.7), and Ozone (r=0.7).

104. *Chromium Fumes*. Fumes generated during high temperature processes involving chromium (Cr)-containing alloys or ores. Exposure can occur during foundry work, welding operations, flame cutting, etc. High exposure has been coded to stainless steel or high chromium (Cr) alloy steel welders, especially to those using manual metal arc welding techniques. Main occupations: welders and flame cutters; industrial equipment mechanics; pipefitters and plumbers. Automatics: chromium compounds; chromium (VI) compounds. Lifetime prevalence: 3.2%. Top three positively correlated chemicals: Nickel Fumes (r=1.0), Nickel Compounds (r=0.6), and Chromium (VI) Comp. (r=0.6).

105. *Manganese Fumes*. Fumes generated during high temperature processes involving manganese (Mn) or manganese-containing alloys. Since all commercial steel contains some manganese which has been introduced in the process of deoxidizing and desulphurizing and to build the strength of the steel, almost all welders of mild and stainless steel have been coded for

manganese fume exposures. Main occupations: welders and flame cutters; pipefitters and plumbers; foundry workers. Automatics: manganese compounds. Lifetime prevalence: 5.2%. Top three positively correlated chemicals: Manganese Compounds (r=0.9), Iron Fumes (r=0.7), and Calcium Oxide Fumes (r=0.7).

106. *Iron Fumes*. Fumes generated during high temperature processes involving iron (Fe) or iron-containing alloys. Exposure to iron fumes occurred during smelting, foundry work, welding and flame cutting operations. Main occupations: welders and flame cutters; motor vehicle mechanics; pipefitters and plumbers. Automatics: iron compounds. Lifetime prevalence: 8.8%. Top three positively correlated chemicals: Calcium Oxide Fumes (r=0.8), Manganese Fumes (r=0.7), and Manganese Compounds (r=0.7).

107. *Nickel Fumes*. Fumes generated during high temperature processes involving nickel (Ni)-containing alloys or ores. The use of high alloyed steels containing a higher proportion of nickel has been increasing in the chemical and aircraft industries resulting in increased exposure to nickel containing aerosols. Exposure occurred mainly during smelting, casting, welding and flame cutting operations. Main occupations: welders and flame cutters; industrial equipment mechanics; pipefitters and plumbers. Automatics: nickel compounds. Lifetime prevalence: 3.2%. Top three positively correlated chemicals: Chromium Fumes (r=1.0), Nickel Compounds (r=0.7), and Chromium (VI) Comp. (r=0.6).

108. *Copper Fumes*. Fumes generated during high temperature processes involving copper (Cu), copper-containing alloys or ores. Exposure to copper fumes generally occurs during welding, casting, or smelting and refining operations. Main occupations: welders and flame cutters; pipefitters and plumbers; foundry molders and coremakers. Automatics: copper compounds. Lifetime prevalence: 3.5%. Top three positively correlated chemicals: Zinc Fumes (r=0.6), Copper Compounds (r=0.5), and Tin Fumes (r=0.5).

109. *Zinc Fumes*. Fumes generated during high temperature processes involving zinc (Zn)-containing alloys or ores. Exposure can occur during foundry and galvanizing operations, brass, bronze and babbitt making and during welding of zinc-containing alloys such as those used in roofing material, pipelines, appliances, and other galvanized materials. Main occupations: welders and flame cutters; pipefitters and plumbers; construction electricians. Automatics: zinc compounds. Lifetime prevalence: 3.3%. Top three positively correlated chemicals: Copper Fumes (r=0.6), Zinc Compounds (r=0.6), and Tin Fumes (r=0.5).

110. *Silver Fumes*. Fumes generated during high temperature processes involving silver (Ag), silver-containing alloys or ores. Exposure occurred mainly in foundry, flame cutting, welding, soldering and jewelry making occupations. Main occupations: jewellers; tools and die makers; welders and flame cutters. Automatics: silver compounds. Lifetime prevalence: 1.5%. Top three positively correlated chemicals: Silver Compounds (r=0.7), Borates (r=0.5), and Copper Fumes (r=0.4).

111. *Tin Fumes*. Fumes generated during high temperature processes involving tin and tin (Sn)-containing alloys or ores. Tin melts at a relatively low temperature (232°C). It is used extensively in solder alloys. Tin fumes have been coded to workers in plumbing and pipe-fitting occupations where gas welding or torches are used, but they have not been coded to workers using electric soldering irons which operate at a much lower temperature. Main occupations: pipefitters and plumbers; welders and flame cutters; motor vehicle mechanics. Automatics: tin compounds. Lifetime prevalence: 4.3%. Top three positively correlated chemicals: Lead Fumes (r=0.8), Tin Compounds (r=0.7), and Soldering Fumes (r=0.5).

112. *Lead Fumes*. Fumes generated during high temperature processes involving lead (Pb), lead-containing alloys or lead-containing ores. During roasting of ores a substantial amount of lead is

released into the environment. Exposure has been especially prominent in certain foundry operations, in the soldering of tin cans and radiators, in the recycling of battery plates and babbitt metal, in the manufacturing of shots and bullets, in the spraying of molten lead alloys and in the casting of type metals. Main occupations: pipefitters and plumbers; welders and flame cutters; sheet metal workers. Automatics: lead compounds. Lifetime prevalence: 3.9%. Top three positively correlated chemicals: Tin Fumes (r=0.8), Tin Compounds (r=0.5), and Soldering Fumes (r=0.5).

113. *Other Pyrolysis Fumes*. A mixture of gases, fumes and particulates of variable composition generated by the heating or burning of organic substances. Included are those fumes which did not fit in any of the other pyrolysis or combustion product categories on our list. Examples are the pyrolysis of paint during heat stripping or welding of coated surfaces, or fumes generated during welding of oil-covered surfaces. Main occupations: welders and flame cutters; tool and dye makers; pipefitters and plumbers. Automatics: PAHs from any source; PAHs from other sources. Lifetime prevalence: 16.8%. Top three positively correlated chemicals: PAH (Other) (r=0.9), Metal Oxide Fumes (r=0.6), and Iron Compounds (r=0.5).

114. *Cooking Fumes*. A mixture of volatile substances of variable composition resulting from the thermal degradation of fats and other food constitutents. Significant quantities of aliphatic aldehydes (formaldehyde and acrolein) have been measured. The temperature and method used for cooking (deep-frying, roasting, charcoal broiling), the type of fat involved and the number of times it has previously been heated can influence the level of contaminants present in the resulting fumes. Main occupations: chefs and cooks; bakers; restaurant managers. Automatics: aliphatic aldehydes. Lifetime prevalence: 6.3%. Top three positively correlated chemicals: Aliphatic Aldehydes (r=0.3), Natural Gas Comb.Prod. (r=0.3), and Flour Dust (r=0.3).

115. *Gasoline Engine Emissions*. Emissions of internal combustion engines running on leaded or unleaded gasoline (automobiles, aircraft, lawnmovers, motorboats, chainsaws). Main occupations: truck, taxi and car (driver-salesmen) drivers; motor vehicle mechanics; woodcutters. Automatics: alcanes (C5-C17); benzo(a)pyrene; carbon monoxide; lead compounds; nitrogen oxides; PAHs from any source; PAHs from petroleum; sulphur dioxide. Lifetime prevalence: 41.5%. Top three positively correlated chemicals: Lead Compounds (r=0.8), Carbon Monoxide (r=0.7), and PAH (Petroleum) (r=0.7).

116. *Coal Combustion Products*. A mixture of gases and particulates generated when coal is used as a heat or energy source. Includes variable amounts of particulates such as carbon, silica, alumina and iron oxides as well as gases such as aldehydes, carbon monoxide, nitrogen oxides, hydrocarbons and sulphur oxides. Coal combustion has been widespread in certain industries and was also widely used for domestic purposes until the 1950's. Main occupations: railway transport workers; construction workers; stationary engineers. Automatics: benzo(a)pyrene; carbon monoxide; nitrogen oxides; PAHs from any source; PAHs from other sources; sulphur dioxide. Lifetime prevalence: 4.5%. Top three positively correlated chemicals: PAH (Coal) (r=0.8), Coal Dust (r=0.5), and Benzo(a)pyrene (r=0.4).

117. *Diesel Engine Emissions*. Emissions of internal combustion engines running on diesel fuels. Engines operating on diesel fuels are used in mines and quarries, railways, buses, trucks etc. Although many workers exposed to gasoline engine emissions would also be exposed to smaller amounts of diesel exhaust, a separate exposure to diesel emissions was coded only when the worker was exposed to higher than environmental background levels. Main occupations: truck drivers, bus drivers, heavy machinery operators. Automatics: alkanes (C5-C17); benzo(a)pyrene; carbon monoxide; nitrogen oxides; PAHs from any source; PAHs from petroleum; soot; sulphur dioxide. Lifetime prevalence: 15.2%. Top three positively correlated chemicals: Diesel Oil (r=0.3), PAH (Petroleum) (r=0.3), and PAH (Any) (r=0.3).

118. *Liquid Fuel Combustion Products*. A mixture of gases and particulates generated when liquid fuel is used as a heat or energy source. Includes variable amounts of gases such as carbon monoxide, nitrogen oxides, sulphur dioxide. Liquid fuel is mainly used as a heating fuel in domestic, commercial and industrial heating installations. Main occupations: stationary engineers; pipefitters and plubmers; construction workers. Automatics: carbon monoxide; nitrogen oxides; sulphur dioxide. Lifetime prevalence: 6.7%. Top three positively correlated chemicals: Sulphur Dioxide (r=0.4), Heating Oil (r=0.4), and Vanadium Compounds (r=0.4).

119. *Wood Combustion Products*. A mixture of gases and particulates generated when wood is used as a heat or energy source. Includes variable amounts of gases such as aldehydes and carbon monoxide, benzo(a)pyrene and other PAHs . Main occupations: farmers; firefighters; chefs and cooks. Automatics: aliphatic aldehydes; benzo(a)pyrene; carbon monoxide; PAHs from any source; PAHs from wood. Lifetime prevalence: 4.2%. Top three positively correlated chemicals: PAH (Wood) (r=1.0), Phosgene (r=0.4), and Benzo(a)pyrene (r=0.3).

120. *Natural Gas Combustion Products*. A mixture of gases generated when natural gas is used as a heat or energy source. Contains substantial amounts of nitrogen oxides but unlike most other combustion products, little carbon monoxide. Natural gas has been widely available in Montreal since 1957 and was used extensively for cooking purposes. Main occupations: food and beverage workers; sheet metal workers; stationary engineers. Automatics: nitrogen oxides. Lifetime prevalence: 3.1%. Top three positively correlated chemicals: Natural Gas (r=0.7), Methane (r=0.5), and Alkanes (C1-C4) (r=0.4).

121. *Jet Fuel Engine Emissions*. The combustion products of the kerosene-type fuel used mainly to power civil and some military aircraft. It contains some of the same constituents as gasoline exhaust although concentrations are reportedly smaller in jet exhaust. Main occupations: air transport workers; aircraft mechanics. Automatics: alkanes (C5-C17); benzo(a)pyrene; carbon monoxide; nitrogen oxides; PAHs from petroleum; PAHs from any source; sulphur dioxide. Lifetime prevalence: 0.8%. Top three positively correlated chemicals: Jet Fuel (r=0.5), Aviation Gasoline (r=0.3), and Chlorinated Alkenes (r=0.2).

122. *Propane Engine Emissions*. Emissions of internal combustion engines running on propane. Includes variable amounts of gases such as carbon monoxide, nitrogen oxides, and some hydrocarbons resulting from incomplete combustion. Potential exposure to propane engine emissions are coded mainly to propane powered lift trucks operators. Main occupations: material handlers; stevedores and freight handlers; shippers and receivers. Automatics: alkanes (C1-C4); carbon monoxide; nitrogen oxides. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Propane (r=0.2), Alkanes (C1-C4) (r=0.2), and PAH (Petroleum) (r=0.1).

123. *Plastics Pyrolysis Products*. A mixture of gases, fumes and soot resulting from the thermal degradation of plastic products. Individual constituents vary depending on the type of plastic and the temperature involved. Carbon monoxide, carbon dioxide, methane, aliphatic and aromatic hydrocarbons are the main gases. The fumes often contain products resulting from an incomplete combustion, such as aldehydes, fatty acids and oligomers, etc. Main occupations: firefighters; dental prosthesis makers; plastic molders. Automatics: PAHs from any source; PAHs from other sources. Lifetime prevalence: 2.9%. Top three positively correlated chemicals: Styrene (r=0.4), Hydrogen Cyanide (r=0.4), and Phosgene (r=0.4).

124. *Rubber Pyrolysis Products*. A mixture of gases, fumes and soot resulting from the thermal degradation of rubber (both natural and synthetic) or rubber products. The composition of the resulting fume varies greatly and depends mainly on the type of rubber, the presence of chemical additives and on the temperature of the process. Typically, these mixtures could contain amines, N-nitrosamines, organic sulphides, carbon disulphide and PAHs. Main occupations: firefighters; rubber bonding workers; rubber mixers. Automatics: PAHs from any source; PAHs from other

sources. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Phosgene (r=0.5), Hydrogen Cyanide (r=0.4), and Plastics Pyrol.Prod. (r=0.4).

125. *Propane Combustion Products*. Propane gas is a convenient combustion fuel because it is marketed in bottles, as opposed to natural gas which is delivered though permanent gas lines. Exposure to the combustion products of propane occurs mainly because of its use as a cooking fuel (especially where natural gas is unavailable, e.g., trains) a welding torch fuel (especially used in plumbing, pipe-fitting and jewelry work) and as a fuel for softening roofing asphalt. Main occupations: pipefitters and plumbers; chefs and cooks; restaurant managers. Automatics: nitrogen oxides. Lifetime prevalence: 3.1%. Top three positively correlated chemicals: Propane (r=0.9), Alkanes (C1-C4) (r=0.6), and Lead Fumes (r=0.3).

126. *Inorganic Acid Solutions*. Inorganic acids are high volume chemicals used extensively in chemical process industries (fertilizers, soap, rayon, film and explosives manufacturing, etc.). Solutions of hydrochloric, sulphuric and nitric acids are the main substances included in this category. They have been used in batteries, as metal cleaners, as chemical reagents in laboratories and in the pharmaceutical industry. Main occupations: service station attendants; auto mechanics; metal platers. Lifetime prevalence: 13.0%. Top three positively correlated chemicals: Sulphuric Acid (r=0.6), Hydrogen Chloride (r=0.5), and Hydraulic Fluid (r=0.3).

127. *Alkali, Caustic Solutions*. Sodium and potassium hydroxide, known respectively as caustic soda and caustic potash, are the main chemicals in this category. However, exposure to other alkaline solutions (e.g., sodium carbonate solutions) or alkaline solutions of unknown composition were also included here. Caustics have been used in the manufacture of rayon, mercerized cotton, soap, paper, explosives and dyestuffs. They have also been used in textile scouring and cleaning baths, for the etching of aluminium, for tin plating, for water softening, as oven cleaners, as drain openers and in laundering and bleaching. Main occupations: stationary engineers; pipefitters and plumbers; janitors. Lifetime prevalence: 7.0%. Top three positively correlated chemicals: Sodium Carbonate (r=0.4), Bleaches (r=0.3), and Inorg.Acid Solutions (r=0.3).

128. *Javel Water*. A clear solution containing sodium hypochlorite and sodium chloride with a strong irritating odor which is known by several trade names such as Javex, Chlorosol and Chlorox. It is widely used as a household bleach and disinfectant and as a bleaching agent in the textile industry. Main occupations: janitors; launderers; butchers. Automatics: hypochlorites. Lifetime prevalence: 5.3%. Top three positively correlated chemicals: Hypochlorites (r=1.0), Biocides (r=0.5), and Cleaning Agents (r=0.4).

129. *Plating Solutions*. Includes all the electrolytes used for electroplating processes, whether the base product is metal or plastic. These are aqueous solutions containing, among other constituents, the salt of the metal being plated. Several inorganic acids are also used routinely, namely, boric, hydrochloric, hydrofluoric and sulphuric. Alkaline solutions are based primarily on sodium or potassium hydroxide. Main occupations: electroplaters; electrotypers; machine tool operators. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Hydrogen (r=0.4), Cyanides (r=0.4), and Hydrogen Cyanide (r=0.4).

130. *Nitric Acid*. A reddish fuming liquid usually marketed in aqueous solutions. The main uses of nitric acid are in the production of fertilizers and explosives. It has also been used in metal degreasing, electroplating, and as a reagent in chemical laboratories. Main occupations: electroplaters; jewellers; engravers. Automatics: inorganic acid solutions. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Plating Solutions (r=0.4), Inorg.Acid Solutions (r=0.3), and Cadmium Compounds (r=0.3).

131. *Phosphoric Acid.* A colorless, syrupy liquid usually sold as an aqueous solutions containing between 10 and 90% acid. It is mostly converted into calcium or ammonium phosphates for fertilizers, but can also be used to etch metals for better paint adhesion. Phosphoric acid has also been used in the textile, rubber and food industries. Main occupations: printing press (offset) operators; motor vehicle refinishers; motor vehicle mechanics. Automatics: inorganic acid solutions. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Inorg.Acid Solutions (r=0.3), Inks (r=0.3), and Aliphatic Esters (r=0.2).

132. *Sulphuric Acid.* An oily, highly corrosive liquid made by burning sulphur to the dioxide, oxidizing to the trioxide and reacting with steam. Produced industrially for over 200 years, this is an important raw material in the manufacture of fertilizers, rayon, and soap and is also commonly used in chemistry laboratories and in the pharmaceutical industry. It has also been used in the pickling and cleaning of metals, as an electrolyte in batteries, and in the purification of petroleum products. Main occupations: motor vehicle mechanics; sheet metal workers; tool and die makers. Automatics: inorganic acid solutions. Lifetime prevalence: 9.8%. Top three positively correlated chemicals: Inorg.Acid Solutions (r=0.6), Hydraulic Fluid (r=0.4), and Sulphur Dioxide (r=0.4).

133. *Methanol.* Also known as methyl alcohol, this chemical is the first member of a homologous series of monohydric aliphatic alcohols. It can be obtained by destructive distillation of wood at about 350°C, but is now mainly synthesized from carbon monoxide. Methanol has been used as the starting material in the manufacture of many chemical products (e.g., formaldehyde), as a solvent in inks, paints and varnishes and for fur cleaning. It is also used in antifreeze mixtures and as an additive for aircraft fuel injection fluids. Main occupations: motor vehicle mechanics; construction painters; service station attendants. Automatics: aliphatic alcohols. Lifetime prevalence: 4.9%. Top three positively correlated chemicals: Aliphatic Alcohols (r=0.7), Ethylene Glycol (r=0.6), and Hydraulic Fluid (r=0.4).

134. *Ethanol.* A colorless, flammable, volatile liquid with a pleasant odor but a burning taste, produced by fermentation of carbohydrates or synthetically from acetylene or ethylene. Also known as ethyl alcohol, it is used in variety of alcoholic beverages. Industrial or denatured alcohol is used as a solvent in the manufacturing of drugs, plastics, lacquers, polishes, perfumes, cosmetics and rubber accelerators. Main occupations: barbers and hairdressers; brewery workers; biologists. Automatics: aliphatic alcohols. Lifetime prevalence: 1.6%. Top three positively correlated chemicals: Aliphatic Alcohols (r=0.4), Cosmetic Talc (r=0.3), and Hair Dust (r=0.3).

135. *Ethylene Glycol.* A colorless syrupy liquid produced from ethylene dichloride or ethylene chlorohydrin and lime. It is used mainly as anti-freeze for automobile-engine cooling systems but has also been used in the production of explosives (nitrated), in cellophane, hydraulic fluids, adhesives and radio condenser pastes. Main occupations: motor vehicle mechanics; construction painters; service station attendants. Lifetime prevalence: 4.1%. Top three positively correlated chemicals: Methanol (r=0.6), Hydraulic Fluid (r=0.5), and Aliphatic Alcohols (r=0.5).

136. *Isopropanol.* A colorless, flammable, mobile liquid, produced by the hydration of propylene from cracked gases. It has been used mainly in the manufacture of acetone, but is also used in extraction processes, as a solvent (chiefly for oils, perfumes and synthetic resins), in liniments, skin lotions, cosmetics and pharmaceuticals. It has been used in rubbing alcohols and as an antistalling agent in winter grade motor fuels. Main occupations: motor vehicle mechanics; barbers and hairdressers; printing press (offset) operators. Automatics: aliphatic alcohols. Lifetime prevalence: 4.1%. Top three positively correlated chemicals: Aliphatic Alcohols (r=0.7), Methanol (r=0.4), and Glycol Ethers (r=0.4).

137. *Acetic Acid.* An organic acid with a pungent odor and sour taste, it is manufactured commercially by the oxidation of ethyl alcohol. It is the principal ingredient of vinegar. In industry it has been used as a raw material for a variety of chemical syntheses, as a mild acid in textile dyeing and printing, and as a solvent for insecticides. Main occupations: photographic processors; painters; photoengravers. Lifetime prevalence: 3.6%. Top three positively correlated chemicals: Arsenic Compounds (r=0.4), Bleaches (r=0.3), and Pesticides (r=0.2).

138. *Carbon Tetrachloride.* A nonflammable heavy, colorless liquid obtained by the chlorination of carbon disulphide. It is an exceptionaly good solvent. Before the 1970s it was extensively used as a degreaser in metal fabricating and in the textile dry cleaning industry. It was also used as a solvent in household spot removing products and in fire extinguishers. Now it is mainly used in the production of fluorocarbons and chlorinated rubbers and as a grain fumigant. Main occupations: firefighters; metal machinists; electricians. Automatics: chlorinated alkanes. Lifetime prevalence: 4.4%. Top three positively correlated chemicals: Chlorinated Alkanes (r=0.6), Phosgene (r=0.4), and Trichloroethylene (r=0.3).

139. *Methylene Chloride.* Also known as dichloromethane, this nonflammable colorless liquid is prepared by the chlorination of chloromethane or methane. It has outstanding solvent properties. It has been used as a paint remover and solvent degreaser, in aerosol formulations, and as a solvent in food and drug processing. Main occupations: construction painters; paint mixers; cabinet and wood furniture makers. Automatics: chlorinated alkanes. Lifetime prevalence: 2.1%. Top three positively correlated chemicals: Chlorinated Alkanes (r=0.4), Titanium Dioxide (r=0.4), and Alkyds (r=0.4).

140. *1,1,1.-Trichlorethane.* Also known as methyl chloroform, this colorless liquid has exceptional solvent and nonflammable properties making it useful as a cleaning solvent for electric motors, generators, and many other electrical and electronic apparatus. It has replaced carbon tetrachloride in metal degreasing (which was banned because of the toxicity associated with this product). It has also been used as a chemical intermediate in the production of vinylidene chloride. Main occupations: electricians; industrial equipment mechanics; rail transport equipment mechanics. Automatics: chlorinated alkanes. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Chlorinated Alkanes (r=0.3), Trichloroethylene (r=0.3), and Chlorinated Alkenes (r=0.3).

141. *Trichloroethylene.* A colorless liquid prepared from 1,2-dichloroethane. Also known as Tri-clene®, it has been widely used as a solvent in vapor degreasing since the early 1930s; it provides economical cleaning of greases, tars, oils and fats from metal parts. Other applications have included decaffeinating coffee (as an extraction solvent) and as a solvent for adhesives and lubricants. Main occupations: machinists; aircraft mechanics; industrial equipment mechanics. Automatics: chlorinated alkenes. Lifetime prevalence: 2.8%. Top three positively correlated chemicals: Chlorinated Alkenes (r=0.8), Carbon Tetrachloride (r=0.3), and 1,1,1.-Trichlorethane (r=0.3).

142. *Perchloroethylene.* A stable colorless liquid, also known as tetrachloroethylene. It is nonflammable and has exceptionally good solvent properties. As a solvent it is used in both cold cleaning and vapor degreasing of metals and is the solvent of choice in the textile dry cleaning industry. It has also been used as a chemical intermediate for the production of fluorocarbons, and to a lesser extent as a heat-exchange fluid, and as a drug against hook worms. Main occupations: dry cleaners; aircraft mechanics; industrial equipment mechanics. Automatics: chlorinated alkenes. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Chlorinated Alkenes (r=0.5), Carbon Tetrachloride (r=0.2), and Trichloroethylene (r=0.2).

143. *Acetone.* A family of organic compounds represented by general formula RCOR', containing a carbonyl group (=CO) linked to two carbon atoms. Acetone (which is also coded

separately) and methyl ethyl ketone (MEK) are widely used in industry as solvents for resins (nitrocellulose, acrylic and epoxies) and for synthetic adhesives. Main occupations: painters; carpenters; motor vehicle mechanics. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Aliphatic Ketones (r=0.6), Aliphatic Esters (r=0.3), and Cellulose Nitrate (r=0.2).

144. *Benzene.* A clear, volatile liquid, derived from coal or petroleum. Industrial benzene which may contain impureties is often known as benzol. It is used as a reagent (chemical and pharmaceutical industries), a solvent (rubber and adhesive industries) and as a constituent of motor fuels. Exposure to benzene was often the result of exposure to solvent mixtures such as pre-1970 mineral spirits. Main occupations: motor vehicle mechanics; service station attendants; shoemakers. Automatics: MAHs. Lifetime prevalence: 17.5%. Top three positively correlated chemicals: Xylene (r=0.7), Leaded Gasoline (r=0.6), and MAH (r=0.6).

145. *Toluene.* A liquid derived from coal or petroleum. Industrial toluene which may contain impureties is often known as toluol. It is used as a solvent (in paints, inks and rubber adhesives), in the production of explosives, dyestuffs and many other chemicals, and in gasoline blending. Exposure to toluene was often the result of exposure to solvent mixtures such as pre-1970 mineral spirits. Main occupations: motor vehicle mechanics; motor vehicle refinishers; carpenters. Automatics: MAHs. Lifetime prevalence: 13.9%. Top three positively correlated chemicals: Xylene (r=0.8), Benzene (r=0.6), and MAH (r=0.6).

146. *Xylene.* A clear liquid produced both from petroleum and coal tar and marketed principally as a mixture of ortho, meta and para isomers. Itis generally referred to as mixed xylenes. Industrial grades of xylene which may contain impureties are often called xylol. Xylenes are used as solvents in lacquers, varnishes, inks, dyes and adhesives and as components in gasoline. Exposure to xylene was often the result of exposure to solvent mixtures such as pre-1970 mineral spirits. Main occupations: motor vehicle mechanics; motor vehicle refinishers; shoemakers. Automatics: MAHs. Lifetime prevalence: 11.2%. Top three positively correlated chemicals: Toluene (r=0.8), Benzene (r=0.7), and MAH (r=0.5).

147. *Styrene.* A colorless viscous liquid, produced mainly by dehydrogenation of ethylbenzene. It is mainly used in the production of plastic resins and synthetic rubbers such as polystyrene plastics and foams, acrylonitrile-butadiene-styrene (ABS), and styrene-acrylonitrile (SAN) resins, and styrene-butadiene rubber (SBR). Occupational exposure has occurred in plants producing the monomer, in polymerization plants and during fabrication of plastic products from unsaturated polyesters dissolved in styrene. Main occupations: motor vehicle mechanics; motor vehicle refinishers; plastic molders. Automatics: MAHs. Lifetime prevalence: 1.8%. Top three positively correlated chemicals: Isocyanates (r=0.6), Phosgene (r=0.5), and Plastics Pyrol.Prod. (r=0.4).

148. *Phenol.* A white crystalline material derived from coal tar. This aromatic alcohol is mainly used in the chemical industry to manufacture phenol-formaldehyde resins, bisphenol-A (used to manufacture epoxy and polycarbonate resins) and various other chemicals. It has also been used as a wound disinfectant and a bactericide. Low-level exposures can also result from thermal degradation of phenol-formaldehyde resins (in foundries where these resins are binders for sand molds and cores or in electrical motor varnishing operations where phenol-formaldehyde is used as an electrical insulating varnish). Main occupations: electric motor repairmen; foundry workers; brewery workers. Automatics: aromatic alcohols. Lifetime prevalence: 1.4%. Top three positively correlated chemicals: Aromatic Alcohols (r=0.7), Phenol-Formald. (r=0.4), and Graphite Dust (r=0.3).

149. *Animal and Vegetable Glues.* Natural glues that include commercial gelatine (hydrolyzed animal collagen, hides, bones, etc.), casein (casein with lime), soybean adhesives (soybean flour

dispersed in an alkaline solution), dextrines (hydrolysed starches), and mucilages. These adhesives have been used for wood and paper products but their use has declined somewhat since the advent of synthetic adhesives. Main occupations: painters and paper hangers; plasterers; cabinet and wood furniture makers. Lifetime prevalence: 3.7%. Top three positively correlated chemicals: Synthetic Adhesives (r=0.3), Cellulose (r=0.2), and Wood Varnishes, Stains (r=0.2).

150. *Turpentine.* An oil obtained by steam distillation of the resin which exudes when various conifer trees are cut. The exact composition of this mixture of terpenes varies according to the country and tree of origin. It has long been recognized as an important solvent and thinner in the paint industry and as a source of resins; it has also been used in the manufacture of linoleum, soap and inks. Main occupations: painters; carpenters; ship workers. Lifetime prevalence: 5.6%. Top three positively correlated chemicals: Linseed Oil (r=0.6), Basic Lead Carb. (r=0.5), and Other Paints,Varnishes (r=0.5).

151. *Linseed Oil.* A mixture of the glycerides of linolic, linoleic, stearic and palmitic acid. It is a golden yellow, amber or brown liquid that is classified as a drying oil; it has been used for years as a binder in paints and in foundry molds. Main occupations: painters, foundry molders and core-makers; cabinet and wood furniture makers. Lifetime prevalence: 4.8%. Top three positively correlated chemicals: Turpentine (r=0.6), Basic Lead Carb. (r=0.6), and Zinc Oxide (r=0.4).

152. *Synthetic Adhesives.* Includes all adhesives based on synthetic resins and rubbers such as formaldehyde resins, epoxy resins, polyvinyl acetate resins and hot melts. Many of these adhesives contain organic solvents. Adhesives are used in many industries, particularly cabinet making and in the furniture and shoe industries. Main occupations: shoemakers; carpenters; cabinet and wood furniture makers. Lifetime prevalence: 14.7%. Top three positively correlated chemicals: Polychloroprene (r=0.4), Urea-Formald. (r=0.4), and Solvents (r=0.4).

153. *Solvents.* Organic liquids used as paint thinners, spot removers, dry cleaning agents, diluents, degreasers, chemical reagents, liquid extraction agents, and for many other purposes. Among the first organic liquids used for this purpose were turpentine, benzene, gasoline and naphtha. More recently, non-flameable chlorinated hydrocarbons came into wider use. Main occupations: motor vehicle mechanics; painters; metal machinists. Lifetime prevalence: 39.1%. Top three positively correlated chemicals: Alkanes (C5-C17) (r=0.6), MAH (r=0.6), and Mineral Spirits+BTX (r=0.5).

154. *Waxes, Polishes.* Includes waxes and polishes for floors, automobiles, leather and furniture. These may contain a variety of substances of animal and vegetable origin such as fatty acids in combination with higher alcohols, petroleum distillates (kerosene, mineral spirits, paraffin waxes), abrasives, and perfumes. Main occupations: janitors; shoemakers; firefighters. Lifetime prevalence: 5.8%. Top three positively correlated chemicals: Cleaning Agents (r=0.4), Biocides (r=0.4), and Javel Water (r=0.3).

155. *Leaded Gasoline.* A mixture of hydrocarbons in the C4 to C12 range produced from petroleum and consisting mainly of straight-chain paraffins which boil within the temperature range of about 30°C to 200°C. It is used as a fuel for automobiles, marine engines and other small engines. In the years covered by this study most gasoline was blended with lead alkyls to increase its octane number. Main occupations: garage mechanics and repairmen; service station attendants; farmers. Automatics: alkanes (C5-C17); benzene; lead compounds; MAHs; toluene; xylene. Lifetime prevalence: 11.9%. Top three positively correlated chemicals: Benzene (r=0.6), Alkanes (C5-C17) (r=0.5), and MAH (r=0.4).

156. *Kerosene.* A petroleum fraction boiling between approximately 180°C and 320°C. It usually consists of a mixture of hydrocarbons containing 10 to 16 carbons per molecule. It has

been widely used as an illuminant, as a cleaning solvent and in insecticides. It is chemically similar to jet fuel and to some heating oils. Main occupations: woodcutters; farmers; printshop workers. Automatics: alkanes (C5-C17); MAHs. Lifetime prevalence: 5.4%. Top three positively correlated chemicals: Alkanes (C5-C17) (r=0.3), Wood Dust (r=0.2), and Leaded Gasoline (r=0.2).

157. *Diesel Oil.* Complex combination of hydrocarbons produced by the distillation of crude oil. It consists of hydrocarbons having carbon numbers predominantly in the range C9-C20 and boiling in the range of approximately 163°C to 375°C. This category encompasses all grades of diesel fuel from light automotive fuels up to heavy marine fuels. Main occupations: motor vehicle mechanics; other mechanics; truck drivers. Automatics: alkanes (C5-C17); PAHs from any source; PAHs from petroleum; MAHs. Lifetime prevalence: 4.1%. Top three positively correlated chemicals: Diesel Eng.Emissions (r=0.3), Leaded Gasoline (r=0.3), and Alkanes (C5-C17) (r=0.3).

158. *Heating Oil.* Also known as fuel oil, heating oil is a mixture of hydrocarbons derived from crude oil by various refining processes. Its chemical composition is similar to that of diesel oil. Two broad classes of heating oil are included in this category: the lighter distillates used as domestic fuel, and the residuals are mainly used in industrial or commercial installations. In this study, most exposures were to the domestic heating oils. Main occupations: stationary engineers; truck drivers (fuel delivery); construction workers. Automatics: alkanes (C5-C17); PAHs from any source; PAHs from petroleum; MAHs. Lifetime prevalence: 4.4%. Top three positively correlated chemicals: Vanadium Compounds (r=0.5), Liquid Fuel Comb.Prod. (r=0.4), and Alkanes (C5-C17) (r=0.3).

159. *Mineral Spirits.* Refined petroleum solvents with carbon chain lengths of C5-C12 and boiling ranges of 150°C-210°C, used since 1970. Various solvent mixtures known as VM&P naphtas (Varnish makers and painters' naphthas), Stoddart solvent and White spirits are included here but purely aliphatic mixtures such as petroleum ethers, which generally boil at lower temperatures, are excluded. A typical chemical composition for mineral spirits would be: 80-86% saturated hydrocarbons, 1% olefins, 0.1% benzene and 13-19% other aromatics. Main occupations: painters; motor vehicle repairmen; stationary engineers. Automatics: alkanes (C5-C17); MAHs. Lifetime prevalence: 11.6%. Top three positively correlated chemicals: Mineral Spirits+BTX (r=0.5), MAH (r=0.5), and Alkanes (C5-C17) (r=0.5).

160. *Lubricating Oils and Greases.* Lubricants are substances which are intended to reduce friction between surfaces in relative motion. They can be of animal, vegetable or mineral origin and although all three types are included in this category, most subjects were exposed to mineral oil-based products. The term greases applies to solid or semi-solid lubricants. Most lubricants are formulated with a variety of additives. Main occupations: motor vehicle repairmen; machinists; farmers. Automatics: alcanes (C18+); benzo(a)pyrene; lead compounds; MAHs; PAHs from any source; PAHs from petroleum. Lifetime prevalence: 30.3%. Top three positively correlated chemicals: Alkanes (C18+) (r=0.9), PAH (Petroleum) (r=0.5), and PAH (Any) (r=0.5).

161. *Cutting Fluids.* Fluids used in metal cutting, machining and drawing processes to cool, clean and reduce friction on the workpieces. This category encompasses all types of cutting fluids whether of animal, vegetable or mineral origin. Before 1955, straight mineral oils predominated whereas now emulsified cutting fluids (mixtures of straight minerals oils and water in proportions running from 1/10 to 1/50) and synthetic cutting fluids are more widely used. For this reason, this category was further divided into cutting fluids pre 1955 and cutting fluids post 1955. Main occupations: machinists; tool and die makers; pipefitters and plumbers. See cutting fluids pre 1955 and cutting fluids post 1955. Lifetime prevalence: 8.3%. Top three positively

correlated chemicals: Cutting Fluids pre 1955 (r=0.9), Cutting Fluids post 1955 (r=0.8), and Metallic Dust (r=0.5).

162. *Asphalt.* A thermoplastic dark brown to black cementitious substance in which the predominant constituents are bitumens obtained by the processing of petroleum crude oils. Asphalt used for roofing, road surfacing, insulating varnishes, acid resistant paints and similar products may contain earthy materials such as sand or limestone. Main occupations: pavers; roofers; truck drivers (asphalt delivery). Automatics: PAHs from any source; PAH from petroleum. Lifetime prevalence: 3.1%. Top three positively correlated chemicals: Coal Tar and Pitch (r=0.3), Crystalline Silica (r=0.2), and PAH (Coal) (r=0.2).

163. *Coal Tar and Pitch.* By-products of the destructive distillation of coal, coal-tars are complex combinations of hydrocarbons, phenols and heterocyclic compounds, while pitches, which are derived from coal tar, contain PAHs and their methyl and polymethyl derivatives as well as heteronuclear compounds. They have been used for waterproof coatings, for road surfaces and as a chemical feedstock. Petroleum-derived products have gradually replaced such coal-tar based products. Pitch is still used for waterproof marine coatings but its main use is in the manufacture of electrodes. Main occupations: roofers; pipefitters and plumbers; aluminium refinery workers. Automatics: benzo(a)pyrene; PAHs from coal; PAHs from any source. Lifetime prevalence: 2.1%. Top three positively correlated chemicals: PAH (Coal) (r=0.5), Asphalt (r=0.3), and Benzo(a)pyrene (r=0.3).

164. *Creosote.* A brownish to black oily liquid obtained from high temperature carbonization of coal tar. It consists of a mixture of guaïacol, cresol, phenol, pyrol, pyridine and other aromatic compounds. For industrial uses, it is mixed with coal tar and petroleum and used in the preservation of wood for railroad ties, telegraph poles, pilings for piers and blocks for flooring. It has also been used as a harsh disinfectant, an animal dip and as a lubricant for die molds. Main occupations: railway trackmen; roofers; power linemen. Automatics: aromatic alcohols; benzo(a)pyrene; PAHs from any source; PAHs from coal. Lifetime prevalence: 1.0%. Top three positively correlated chemicals: Aromatic Alcohols (r=0.6), PAH (Coal) (r=0.4), and Excavation Dust (r=0.2).

165. *Hydraulic Fluid.* Moving parts of many industrial machines are actuated by hydraulic fluids which are under pressure. Many are mineral-oil based (straight mineral oils or oil-water emulsions) but other oils (i.e., castor oil), solvents (most notably glycols or glycol ethers) and synthetic fluids are also used. Small quantities of various additives are usually present. Main occupations: Motor vehicle mechanics; aircraft mechanics; service station workers. Lifetime prevalence: 3.9%. Top three positively correlated chemicals: Ethylene Glycol (r=0.5), Methanol (r=0.4), and Glycol Ethers (r=0.4).

166. *Other Mineral Oils.* These petroleum-derived oils contain relatively high molecular-weight paraffinic, cycloparaffinic and aromatic hydrocarbons. The composition of these oils is similar to that of some lubricating oils, hydraulic fluids or cutting fluids. This category included textile oils, heat treating oils, rolling oils, drawing oils, rubber oils, forging oils, mold-release oils and the mineral oils used in ink formulations. Main occupations: printshop workers; textile workers; forgers. Automatics: alcanes (C18+); benzo(a)pyrene; MAHs; PAHs from petroleum; PAHs from any source. Lifetime prevalence: 3.6%. Top three positively correlated chemicals: Inks (r=0.4), Carbon Black (r=0.3), and Aromatic Amines (r=0.3).

167. *Jet Fuel.* The two main types of jet fuel, kerosene and wide-cut were included in this category. Kerosene type is a relatively high flash point range petroleum distillate within the kerosene boiling range, typically between 169°C and 235°C. Wide-cut type is a relatively wide boiling range volatile petroleum distillate including both gasoline and kerosene fractions with a range of 61°C to 235°C. Main occupations: aircraft mechanics and repairmen; dockworkers;

refinery pipefitters and plumbers. Automatics: alkanes (C5-C17); MAHs. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Aviation Gasoline (r=0.6), Jet Fuel Eng.Emiss. (r=0.5), and Diesel Oil (r=0.2).

168. *Aviation Gasoline.* Fuel used to power small piston engine aircraft. Although similar in composition to leaded automotive gasoline, aviation gasoline has a higher content of branched alkanes and a more limited boiling range of about 50°C to 170°C. Main occupations: aircraft mechanics and repairmen; military and civil aircraft pilots; refinery pipefitters and plumbers. Automatics: alkanes (C5-C17); benzene; lead compounds; MAHs; toluene; xylene. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Jet Fuel (r=0.6), Jet Fuel Eng.Emiss. (r=0.3), and Benzene (r=0.2).

169. *Mineral Spirits+BTX.* Refined petroleum solvents with carbon chain lengths of C5-C12 and boiling ranges of 150°C-210°C, used before 1970. Before 1970 mineral spirits contained relatively higher amounts of benzene, toluene and xylene due to ignorance of their toxic effects. A typical chemical composition for mineral spirits before this time period would roughly be: 80-86% saturated hydrocarbons, 1% olefins, and 13-19% aromatics (usually containing at least 1% benzene and relatively greater quantities of toluene and xylene than mineral spirits post 1970). Main occupations: motor vehicle repairmen; painters; stationary engineers. Automatics: alkanes (C5-C17); benzene; MAHs; toluene; xylene. Lifetime prevalence: 15.6%. Top three positively correlated chemicals: Alkanes (C5-C17) (r=0.6), MAH (r=0.6), and Mineral Spirits (r=0.5).

170. *Cutting Fluids pre 1955.* Fluids used before 1955 in metal cutting, machining and drawing processes to cool, clean and reduce friction on the workpieces. Prior to 1955, most cutting fluids in use were straight mineral oils. Furthermore, the refining process used in this period did not involve solvent extraction and consequently these oils contained relatively high levels of PAHs. Main occupations: machinists; pipefitters and plumbers; tool and die makers. Automatics: alcanes (C18+); benzo(a)pyrene; cutting fluids; MAHs; PAHs from any source; PAHs from petroleum. Lifetime prevalence: 6.3%. Top three positively correlated chemicals: Cutting Fluids (r=0.9), Cutting Fluids post 1955 (r=0.6), and Metallic Dust (r=0.4).

171. *Cutting Fluids post 1955.* Fluids used since 1955 in metal cutting, machining and drawing processes to cool, clean and reduce friction on the workpieces. Although straight mineral oils are still used, there has been an increased tendency to use aqueous oil emulsions or synthetic cutting fluids. Furthermore, the use of solvent-extracted oils has led to reduced quantities of PAHs in these oils. Main occupations: machinists; pipefitters and plumbers; tool and die makers. Automatics: alcanes (C18+); benzo(a)pyrene; cutting fluids; MAHs; PAHs from any source; PAHs from petroleum. Lifetime prevalence: 5.4%. Top three positively correlated chemicals: Cutting Fluids (r=0.8), Cutting Fluids pre 1955 (r=0.6), and Mild Steel Dust (r=0.4).

172. *Other Paints,Varnishes.* Paints used on surfaces other than metal and varnishes used on surfaces other than wood. A paint is a dispersion of a finely divided pigment in a liquid composed of a resin or binder and a volatile solvent used to cover plaster, wood, gyproc and metal. Trade sale paints such as alkyds, acrylic latexes and caseins (in the past) were included. Main occupations: painters; carpenters; janitors. Lifetime prevalence: 13.2%. Top three positively correlated chemicals: Turpentine (r=0.5), Wood Varnishes, Stains (r=0.4), and Solvents (r=0.4).

173. *Wood Varnishes, Stains.* Varnishes are light-bodied quick drying products that form a glossy or mat finish on application. Oleoresinous varnishes are made of resins in drying oils, mixed with driers and thinning agents such as alcohols, ethers, and naphthas. Varnishes based on alkyds and urethanes have largely replaced them. Stains are varnishes containing enough pigment or dye to alter the appearance of a wood surface. Main occupations: painters;

carpenters; cabinet and wood furniture makers. Lifetime prevalence: 5.3%. Top three positively correlated chemicals: Turpentine (r=0.5), Polyvinyl Acetate (r=0.4), and Alkyds (r=0.4).

174. *Inks*. Colored liquids or pastes used for writing, drawing, marking, and printing. Writing ink usually contains ferrous sulphate and indigo dye with tannic acid. Various formulations of printing inks are available to meet the demands of specific print jobs. Newspaper inks, for instance usually contain carbon black and various mineral oils. Ink pigments may be inorganic (including lead chromates) or organic (including benzidine yellows). Main occupations: typesetters and printers; draughtsmen; business machine repairmen. Lifetime prevalence: 3.5%. Top three positively correlated chemicals: Carbon Black (r=0.4), Other Mineral Oils (r=0.4), and Inorg.Pigments (r=0.3).

175. *Metal Coatings*. Paints and coatings used specifically for metals and made up of combinations of oxidizing alkyds, epoxies, formaldehyde resins, thermosetting acrylics, polyesters and others. A paint is a dispersion of a finely divided pigment in a liquid composed of a resin or binder and a volatile solvent used to cover plaster, wood, gyproc and metal. These coatings may be applied to motor vehicles, structural steel, ships, home applicances, metal furniture, etc. Main occupations: motor vehicle refinishers; construction painters; motor vehicle mechanics. Lifetime prevalence: 7.3%. Top three positively correlated chemicals: Alkyds (r=0.4), Inorg.Pigments (r=0.4), and Extenders (r=0.4).

176. *Cyanides*. Includes hydrogen cyanide (which was also coded separately), sodium, potassium and calcium cyanide and all other salts of hydrocyanic acid. Molten sodium cyanide is used in metallurgy in case-hardening processes and cyanide solutions are widely used in electroplating. Cyanides have also been used to clean brass, copper and other metal surfaces, in the extraction of gold, in photography, and in insecticides. Main occupations: firefighters; electroplaters; jewellers. Lifetime prevalence: 2.0%. Top three positively correlated chemicals: Hydrogen Cyanide (r=0.9), Phosgene (r=0.5), and Plating Solutions (r=0.4).

177. *Fluorides*. Includes exposures to all fluorides (e.g., Na3AlF6, sodium aluminium fluoride, also called cryolite, used as a flux in the production of aluminium, in the fabrication of special glasses, porcelain and in insecticides). Some welding electrode coatings contain a calcium carbonate-calcium fluoride system; this is thermally degraded during welding to silicon hexafluoride which gives rise to hydrogen fluoride in the presence of water. Sodium, potassium and calcium fluorides are also present in the welding environment. Main occupations: mineral ore treaters; welders; aluminium refinery workers. Lifetime prevalence: 3.4%. Top three positively correlated chemicals: Hydrogen Fluoride (r=0.9), Ozone (r=0.5), and Manganese Fumes (r=0.5).

178. *Chromium (VI) Compounds*. Comprises hexavalent chromium (Cr) compounds found in chromium fumes and in various chromium compounds. The most important compounds in this hexavalent state are sodium and potassium dichromate, chromic acid and a number of pigments such as lead chromate, (which was coded separately), zinc chromate and strontium chromate. These compounds have many industrial applications as a consequence of their acidic and oxidant properties and their ability to form strongly colored and insoluble salts. They have been used in the manufacture of important inorganic pigments which are used in paints, artist's colors, glasses and glazes. They have also been used in chrome plating, for corrosion inhibition and in wood preservation, leather tanning, textile dyeing, lithography and chrome plating. Main occupations: metal painters; welders and flame cutters; motor vehicle mechanics. Automatics: chromium compounds. Lifetime prevalence: 8.4%. Top three positively correlated chemicals: Chromium Compounds (r=0.8), Lead Chromate (r=0.6), and Chromium Fumes (r=0.6).

179. *Hypochlorites*. ncludes both sodium and calcium hypochlorites. These compounds decompose easily in water and are used as a source of chlorine for cleaning, bleaching and

sanitizing. A water solution of sodium hypochlorite known as javel water is used extensively in the laundry industry. These bleaching powders have also been used in the textile and paper pulp industries. Main occupations: laundry workers; textile bleachers; janitors. Lifetime prevalence: 5.6%. Top three positively correlated chemicals: Javel Water (r=1.0), Biocides (r=0.5), and Cleaning Agents (r=0.4).

180. *Nitrates*. Inorganic compounds containing one or several -NO3 functional groups. Sodium, potassium and ammonium nitrates are the main substances included here. They have been used in the manufacture of explosives, fertilizers, glass, pyrotechnics, welding fluxes and matches; in steel heat treating, and in meat processing. Main occupations: munitions assemblers; meat packers; pipefitters and plumbers. Lifetime prevalence: 1.0%. Top three positively correlated chemicals: Zinc Dust (r=0.2), Tin Fumes (r=0.2), and Propane Comb.Prod. (r=0.2).

181. *Beryllium Compounds*. Comprises beryllium (Be) fumes, dust from beryllium-containing alloys and ores and all other beryllium-containing substances. Due to the high cost of beryllium, it is not used in engineering or as a construction material. It is increasingly used in the atomic energy industry, and in alloys for anti-spark tools and machinery parts such as bushings and current carrying springs which are subjected to abnormal wear, vibrations, or shocks. Stationary engineers and engine and boiler-room crew in ships have been exposed from coal ashes, which contain beryllium compounds. Main occupations: stationary engineers; metal machinists. Lifetime prevalence: 0.5%. Top three positively correlated chemicals: Cobalt Compounds (r=0.4), Sodium Carbonate (r=0.3), and Ashes (r=0.3).

182. *Magnesium Compounds*. Comprises magnesium (Mg) dust, magnesium fumes (which was also coded separately), magnesium-containing alloys and ores and all other magnesium-containing substances (e.g., magnesium oxide, magnesium chloride, magnesium sulphate and magnesium metasilicates). Magnesium being an active metal, it is rarely used as a construction material. However, due to its lightness it is mainly used as an alloying element. It has been used as a desulphurizer and deoxidizer in the production of copper and nickel based alloys, and in pyrotechnics and signal flares. Magnesium compounds have also been used in insulating cement, pipe insulation and in fertilizers. Magnesite, a mineral composed of magnesium carbonate mixed with some iron carbonate and ferric oxide, is a valued refractory material for crucibles, furnace brick and linings and high temperature electrical insulation. Main occupations: stationary engineers; metal machinists; foundry workers. Lifetime prevalence: 1.6%. Top three positively correlated chemicals: Copper Dust (r=0.2), Nickel Compounds (r=0.2), and Aluminium Alloy Dust (r=0.2).

183. *Aluminium Compounds*. Comprises dust from aluminium (Al)-containing alloys and aluminium fumes (both of which were also coded separately), dust from aluminium-containing ores and all other aluminium-containing substances (some of which were also coded separately, e.g. alumina, alum). The major industrial uses of aluminium have been in the construction, aircraft, and electrical industries. Alumina has been used extensively as an abrasive in a wide variety of machine tools such as fast cutting and grinding wheels and in refractories. Main occupations: metal machinists; carpenters; welders and flame cutters. Lifetime prevalence: 18.8%. Top three positively correlated chemicals: Alumina (r=0.9), Abrasives Dust (r=0.7), and Aluminium Alloy Dust (r=0.5).

184. *Titanium Compounds*. Comprises dust from titanium (Ti)-containing alloys and ores, titanium dioxide dust and fumes (both of which were also coded separately) and all other titanium-containing substances. Titanium metal is obtained commercially from two ores: rutile and ilmenite. It has been used in jet engine components and as an alloying element in steels. Main occupations: construction painters; motor vehicle refinishers; metal machinists. Lifetime

prevalence: 4.6%. Top three positively correlated chemicals: Titanium Dioxide (r=0.9), Alkyds (r=0.6), and Extenders (r=0.6).

185. *Vanadium Compounds*. Comprises vanadium (V) dust, vanadium fumes, and dust from vanadium-containing alloys and ores and all other vanadium-containing substances. Exposure to vanadium compounds has been mainly restricted to mining and milling of vanadium containing ores. Engine and boiler-room workers may also be exposed because the ashes and soot of oil-fired burners contain vanadium pentoxide. Main occupations: stationary engineers. Lifetime prevalence: 1.3%. Top three positively correlated chemicals: Heating Oil (r=0.5), Antimony Compounds (r=0.5), and Sodium Carbonate (r=0.4).

186. *Chromium Compounds*. Comprises chromium (Cr) dust, chromium fumes, chromium (VI) compounds (all of which were also coded separately), dust from chromium-containing alloys (including stainless steel which was coded separately) and ores and all other chromium-containing substances. Many chromium compounds have been used in the manufacture of important inorganic pigments. Main occupations: construction painters; metal machinists; welders and flame cutters. Lifetime prevalence: 12.3%. Top three positively correlated chemicals: Chromium (VI) Comp. (r=0.8), Nickel Compounds (r=0.7), and Stainless Steel Dust (r=0.6).

187. *Manganese Compounds*. Comprises manganese (Mn) dust, manganese fumes (which were also coded separately), dust from manganese-containing alloys and ores and all other manganese-containing substances. Manganese is a silvery-white metal found in a number of minerals, with iron ores and in most rocks. It is used mainly as an alloying element to improve the strength and hardness of steels and and to reduce the oxygen and sulphur naturally present in the iron ore. Some manganese compounds have also been used in the manufacturing of dry cell batteries, as oxidizing agents (e.g., potassium permanganate) in the chemical industry, and as drying agents for linseed oil-based paints. Organo-manganese compounds have also been used as antiknock additives in gasoline. Main occupations: welders and flame cutters; pipefitters and plumbers; metal grinders. Lifetime prevalence: 6.3%. Top three positively correlated chemicals: Manganese Fumes (r=0.9), Iron Fumes (r=0.7), and Calcium Oxide Fumes (r=0.6).

188. *Iron Compounds*. Comprises iron (Fe) dust, iron oxides and iron fumes (all of which were also coded separately), dust from iron-containing alloys (mild and stainless steel were also coded separately), iron-containing ores and all other iron-containing substances. Iron is the most common of the commercial metals and forms a large group of materials known as ferroalloys. Several iron compounds have been used as paint pigments, polishing compounds, and coatings for magnetic tapes while the soluble salts have been used as dyeing mordants, catalysts, fertilizers, in sewage treatments, and in feeds. Main occupations: metal machinists; welders and flame cutters; motor vehicle mechanics. Lifetime prevalence: 24.9%. Top three positively correlated chemicals: Mild Steel Dust (r=0.8), Metallic Dust (r=0.7), and Metal Oxide Fumes (r=0.6).

189. *Cobalt Compounds*. Comprises cobalt (Co) fumes, dust from cobalt-containing alloys and ores and all other cobalt-containing substances (e.g., acetate, oleate, resinate). Most cobalt has been used for high temperature alloys. Cobalt is added to tool steels to increase the strength and hardness of those required to operate at high speed and high temperature. A variety of organic salts of cobalt, such as resinate, oleate and acetate etc. have been used extensively as drying agents for paints, inks and varnishes. Other cobalt compounds have been used in pottery to improve color. Main occupations: tool and die makers; metal machinists; stationary engineers. Lifetime prevalence: 2.0%. Top three positively correlated chemicals: Tungsten Compounds (r=0.7), Beryllium Compounds (r=0.4), and Cutting Fluids post 1955 (r=0.3).

190. *Nickel Compounds*. Comprises nickel (Ni) dust and nickel fumes, (both of which were also coded separately), dust from nickel-containing alloys (including stainless steel, which was coded separately) and ores and all other nickel-containing substances. Most exposures in this study occurred through the use of nickel-based alloys, mainly stainless steel. Other major uses have been in plating where more nickel has been used than any other metal, and as a catalyst in hydrogenation of organic compounds. Main occupations: metal machinists; welders and flame cutters; metal grinders. Lifetime prevalence: 7.0%. Top three positively correlated chemicals: Stainless Steel Dust (r=0.8), Chromium Compounds (r=0.7), and Nickel Fumes (r=0.7).

191. *Copper Compounds*. Comprises copper (Cu) dust and copper fumes (both of which were also coded separately), dust from copper-containing alloys and ores and all other copper-containing substances. Copper compounds have been used in electroplating solutions, insecticides and fungicides; copper-containing alloys have been used in jewelry and silverware. Main occupations: farmers; pipefitters and plumbers; metal machinists. Lifetime prevalence: 12.2%. Top three positively correlated chemicals: Copper Dust (r=0.6), Copper Fumes (r=0.5), and Tin Compounds (r=0.4).

192. *Zinc Compounds*. Comprises zinc (Zn) dust, zinc fumes (both of which were also coded separately), dust from zinc-containing alloys and ores and all other zinc-containing substances (including zinc oxide which was also coded separately). Zinc sulphide, a white powder, has been used as a pigment in paints, for whitening rubber, and for paper coating. Main occupations: construction painters; pipefitters and plumbers; motor vehicle refinishers. Lifetime prevalence: 10.1%. Top three positively correlated chemicals: Zinc Fumes (r=0.6), Zinc Oxide (r=0.5), and Zinc Dust (r=0.5).

193. *Arsenic Compounds*. Comprises arsenic (As) dust, arsenic fumes, dust from arsenic-containing alloys and ores and all other arsenic containing substances (e.g., calcium, sodium and lead arsenate). Although use of the metal is limited to a few applications involving lead-arsenic alloys, several compounds have been widely used as insecticides. The smelting and refining of arsenic-containing ores (e.g., copper smelting and refining) and boiler cleaning (ashes and soot may contain arsenic trioxide) are other potential sources of exposure. Main occupations: general farmers; crop and vegetable farmers; stationary engineers. Lifetime prevalence: 3.5%. Top three positively correlated chemicals: Acetic Acid (r=0.4), Pesticides (r=0.4), and Phosgene (r=0.3).

194. *Silver Compounds*. Comprises silver (Ag) dust, silver fumes (which were also coded separately), dust from silver-containing alloys and ores and all other silver-containing substances. Silver (Ag) is a white metal which occurs in the native state or combined with sulphur and chlorine. Copper, lead and zinc ores frequently contain silver and most of the production of this metal is a by-product of the refining of these metals. Silver and its compounds have been used in photography, electrical applications, jewelry and silverware, brazing alloys, and mirrors. Main occupations: jewellers; tool and die makers; dental prosthesis makers. Lifetime prevalence: 2.7%. Top three positively correlated chemicals: Silver Fumes (r=0.7), Gold Compounds (r=0.5), and Borates (r=0.5).

195. *Cadmium Compounds*. Comprises cadmium (Cd) dust and cadmium fumes (both of which were also coded separately), dust from cadmium-containing alloys and ores and all other cadmium-containing substances (e.g. cadmium sulphide, selenide, nitrate). Cadmium has been widely used in electroplating as a protective coating for iron, steel, and copper. Cadmium compounds such as cadmium sulphide and cadmium selenide are important coloring pigments for plastics, paints, etc. Main occupations: tool and die makers; jewellers; pipefitters and plumbers. Lifetime prevalence: 1.0%. Top three positively correlated chemicals: Silver Fumes (r=0.4), Nitric Acid (r=0.3), and Silver Compounds (r=0.3).

196. *Tin Compounds*. Comprises tin (Sn) dust, tin fumes (both of which were also coded separately), dust from tin-containing alloys and ores and all other tin-containing substances. Organo-tins have been used as plastics stabilizers, biocides and as catalysts. Main occupations: pipefitters and plumbers; motor vehicle mechanics; typesetters and printing press (letterpress) operators. Lifetime prevalence: 8.9%. Top three positively correlated chemicals: Soldering Fumes (r=0.7), Tin Fumes (r=0.7), and Lead Fumes (r=0.5).

197. *Antimony Compounds*. Comprises antimony (Sb) dust, antimony fumes, dust from antimony-containing alloys and ores and all other antimony-containing substances. Antimony itself is a lustrous, silvery blue□white, extremely brittle metal. When alloyed with other metals, it increases hardness, lowers melting points and reduces shrinkage upon freezing. In this study, most exposures were due to antimony-lead alloys used as type metal, storage battery plates, bullets, tank linings, bearing metals, etc. Residual soot or ashes may also contain antimony compounds. Main occupations: stationary engineers; printing press (letterpress) operators; typesetters. Lifetime prevalence: 1.9%. Top three positively correlated chemicals: Vanadium Compounds (r=0.5), Other Mineral Oils (r=0.3), and Inks (r=0.3).

198. *Tungsten Compounds*. Comprises tungsten (W) dust, tungsten fumes, and dust from tungsten-containing alloys and ores and all other tungsten-containing substances. Tungsten (W), a white, heavy metal, is widely distributed in small quantities in nature but is mostly obtained from scheelite, wolframite and a few other ores. It has one of the highest melting points (3400°C) of all metals, a property that renders it very useful for lamp filaments, electric contacts, rocket nozzles and in electronic applications. The most important tungsten compound is the carbide (WC), which has been used in cutting tool bits. Tool steels containing up to 18% of tungsten are generally used in machining steels. Main occupations: metal machinists; tool and die makers; machine tool operators. Lifetime prevalence: 1.4%. Top three positively correlated chemicals: Cobalt Compounds (r=0.7), Cutting Fluids post 1955 (r=0.3), and Silicon Carbide (r=0.3).

199. *Gold Compounds*. Comprises gold (Au) dust, gold fumes (which were also coded separately), dust from gold-containing alloys and ores and all other gold-containing substances. Due to the high cost of gold, it is not used as a construction or engineering material except as a coating on spacecraft for radiation control. Gold has also been used for coinage, ornaments, jewelry, dentistry and (alloyed with platinum and silver) for chemically resistant apparatus. It is nomally alloyed with copper and silver to increase hardness. Main occupations: jewellers; dental prosthesis makers; watch and clock repairmen. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Silver Compounds (r=0.5), Borates (r=0.5), and Silver Fumes (r=0.4).

200. *Mercury Compounds*. Comprises metallic mercury (Hg) (which was also coded separately), dust from mercury-containing amalgams and ores and all other mercury-containing substances. Mercury compounds have been used as agricultural and industrial poisons; some organo-mercuric compounds have been used as antiseptics. Other mercury compounds are or have been used in taxidermy, in carroting rabbit fur for felt hats, in the manufacture of explosives and in paints as antifouling and mildew-proofing agents. Main occupations: construction painters; chemists; paint mixers. Lifetime prevalence: 1.7%. Top three positively correlated chemicals: Poly-Acrylates (r=0.5), Zinc Oxide (r=0.4), and Polyvinyl Acetate (r=0.4).

201. *Lead Compounds*. Comprises lead dust (Pb), lead oxides and lead fumes (all of which were also coded separately), dust from lead-containing alloys and ores and all other lead-containing substances (e.g., lead chromate, basic lead carbonate, which were also coded separately). Inorganic lead compounds have been used in agriculture and as pigments in rubber, plastics and paints. Organic lead compounds, such as tetraethyl or tetramethyl lead, have been used as antiknock compounds in leaded gasolines, while lead alloys have been commonly used as solders. Main occupations: motor vehicle drivers and driver-salesmen; motor vehicle mechanics; painters.

Lifetime prevalence: 46.6%. Top three positively correlated chemicals: Gas Eng.Emissions (r=0.8), Carbon Monoxide (r=0.7), and PAH (Petroleum) (r=0.6).

202. *Alkanes (C18+).* Includes all saturated hydrocarbons having more than 18 carbon atoms, with the general formula CnH2n+2. They are all solids at standard conditions. One mixture of these long-chained hydrocarbons, known as petroleum jelly, is widely used in lubricating oils and greases and for compounding in rubber and resins. Highly refined, it is used in the pharmaceutical industry. Parrafin waxes, which were also coded separately, also fall into this category. Main occupations: motor vehicle mechanics; metal machinists; business and industrial machine mechanics. Lifetime prevalence: 33.6%. Top three positively correlated chemicals: Lubric.Oils & Greases (r=0.9), PAH (Petroleum) (r=0.5), and PAH (Any) (r=0.5).

203. *Alkanes (C1-C4).* Includes the first four similarly structured compounds called paraffins, namely: methane, ethane, propane and butane. They are gaseous at standard temperatures and are used as fuels or as raw materials in the chemical industry. Main occupations: chefs and cooks; pipefitters and plumbers; restaurant managers. Lifetime prevalence: 8.9%. Top three positively correlated chemicals: Methane (r=0.7), Propane (r=0.6), and Propane Comb.Prod. (r=0.6).

204. *Alkanes (C5-C17).* Saturated hydrocarbons, straight or branched-chained with the general formula CnH2n+2, containing between 5 and 17 carbon atoms. They are liquids at standard conditions. They are the main components of petroleum solvents such as petroleum ether, rubber solvent, VM&P naphthas, mineral spirits, Stoddard solvent, kerosene and fuels such as gasoline, jet fuel and heating oils. Main occupations: motor vehicle mechanics; painters; carpenters. Lifetime prevalence: 35.8%. Top three positively correlated chemicals: MAH (r=0.8), Solvents (r=0.6), and Mineral Spirits+BTX (r=0.6).

205. *Aliphatic Alcohols.* Aliphatic hydrocarbon derivatives in which one hydrogen atom is replaced by an -OH group. This includes methanol, ethanol, isopropanol (all of which were also coded separately) and propanol. These alcohols are generally used as solvents in toiletries, pharmaceuticals, and surface coatings, as rubbing compounds, and in fur cleaning. Ethanol is used in alcoholic beverages. Main occupations: motor vehicle mechanics; barbers and hairdressers; service station attendants. Lifetime prevalence: 9.0%. Top three positively correlated chemicals: Methanol (r=0.7), Isopropanol (r=0.7), and Ethylene Glycol (r=0.5).

206. *Aliphatic Aldehydes.* A family of organic compounds represented by general formula RCHO (where R is a hydrogen or an alkyl group). The most important exposures in this group are formaldehyde (which was also coded separately and is widely used in various industries) and acrolein (a pyrolysis product of many organic compounds). Main occupations: chefs and cooks; carpenters; textile workers. Lifetime prevalence: 17.3%. Top three positively correlated chemicals: Formaldehyde (r=0.9), Cooking Fumes (r=0.3), and Plastics Pyrol.Prod. (r=0.3).

207. *Chlorinated Alkanes.* Saturated hydrocarbons in which at least one hydrogen is replaced by a chlorine atom. This replacement increases many desirable properties such as specific gravity and boiling points, and reduces flammability. These materials, e.g., methylene chloride, chloroform, carbon tetrachloride, and 1,1,1-trichloroethane (all of which were also coded separately), are used as solvents for fats and oils, for metal degreasing, for drycleaning of textiles, as refrigerants, in insecticides, and in fire extinguishers. Main occupations: motor vehicle mechanics; construction painters; aircraft mechanics. Lifetime prevalence: 10.0%. Top three positively correlated chemicals: Carbon Tetrachloride (r=0.6), Methylene Chloride (r=0.4), and Chlorinated Alkenes (r=0.4).

208. *Unsaturated Aliphatic Hydrocarbons.* All organic compounds containing only carbon and hydrogen atoms and at least one carbon-carbon double or triple bond, i.e., alkenes and alkynes. Ring compounds are excluded. The main substances included in this category were acetylene,

ethylene (both of which were also coded separately) and propylene. Main occupations: welders and flame cutters; pipefitters and plumbers; motor vehicle mechanics. Lifetime prevalence: 5.2%. Top three positively correlated chemicals: Acetylene (r=0.9), Gas Welding Fumes (r=0.5), and Iron Fumes (r=0.5).

209. *Chlorinated Alkenes*. Unsaturated hydrocarbons in which one or more hydrogens are replaced with chlorine atoms. These relatively nonflammable, organic compounds are used in dry cleaning of textiles and in metal degreasing. Examples are trichloroethylene, perchloroethylene and vinyl chloride (all of which were also coded separately). Main occupations: barbers and hairdressers; metal machinists; aircraft mechanics. Lifetime prevalence: 4.8%. Top three positively correlated chemicals: Trichloroethylene (r=0.8), Perchloroethylene (r=0.5), and Chlorinated Alkanes (r=0.4).

210. *Aliphatic Esters*. Compounds with the general formula RC=OOR' (R and R' are aliphatic groups) produced by reacting an alcohol and an acid. The main substances included here are the formates and the acetates (methyl, ethyl, propyl, butyl, etc.). They have been used mainly as solvents for resins (nitrocellulose), and in the production of varnishes, artificial leather and pharmaceuticals. Main occupations: painters; motor vehicle mechanics; cabinet and wood furniture makers. Lifetime prevalence: 3.3%. Top three positively correlated chemicals: Aliphatic Ketones (r=0.6), Cellulose Nitrate (r=0.6), and Phthalates (r=0.4).

211. *Aliphatic Ketones*. A family of organic compounds represented by general formula RCOR', containing a carbonyl group (=CO) linked to two carbon atoms. Acetone (which is also coded separately) and methyl ethyl ketone (MEK) are widely used in industry as solvents for resins (nitrocellulose, acrylic and epoxies) and for synthetic adhesives. Main occupations: painters; carpenters; motor vehicle mechanics. Lifetime prevalence: 6.0%. Top three positively correlated chemicals: Acetone (r=0.6), Aliphatic Esters (r=0.6), and Cellulose Nitrate (r=0.5).

212. *Fluorocarbons*. Paraffinic compounds in which one or more hydrogen atoms are replaced by fluorine. They may also contain chlorine. They have been used mainly in fire extinguishers, as refrigerants, as cleaning solvents, and as propellants in variety of products ranging from paints and insecticides to cosmetics such as perfumes, hair sprays and deodorants. Main occupations: firefighters; barbers and hairdressers; electric and electronic equipment installers and repairmen. Lifetime prevalence: 2.4%. Top three positively correlated chemicals: Phosgene (r=0.5), Spray Gases (r=0.5), and Hydrogen Cyanide (r=0.4).

213. *Glycol Ethers*. Synthetic organic liquids with sweetish odors and high boiling points generally manufactured by reaction of an epoxide (ethylene, propylene or butylene oxide) with the appropriate alcohol. Miscibility of these ethers with water and organic solvents makes them especially useful as mutual solvents in many oil-water compositions. They have been used as solvents for various resins, lacquers, paints, varnishes, dyes, inks, printing pastes, cleaning agents, liquid soaps, and even cosmetics. They have also been used widely as components of hydraulic fluids and as chemical intermediates. Main occupations: motor vehicle mechanics; janitors; construction painters. Lifetime prevalence: 2.5%. Top three positively correlated chemicals: Hydraulic Fluid (r=0.4), Aliphatic Alcohols (r=0.4), and Ethylene Glycol (r=0.4).

214. *Polycyclic Aromatic Hydrocarbons (Any)*. Polycyclic aromatic hydrocarbons are a group of chemicals made up of three or more benzene rings interlinked in various arrangements. They are naturally present in fossil fuels or can be formed by thermal decomposition of any organic material containing carbon and hydrogen. Because the profile of PAHs produced depends, among other factors, on the source material which gives rise to PAH exposure, four categories corresponding to various classes of source materials were created: PAHs from coal, PAHs from petroleum; PAHs from wood and PAHs from other sources. In addition, exposure to benzo(a)pyrene was coded. The category described here was assigned whenever one of the

specific categories mentioned above was coded. Main occupations: motor vehicle drivers and driver-salesmen; motor vehicle repairmen; machinists. Lifetime prevalence: 63.0%. Top three positively correlated chemicals: PAH (Petroleum) (r=0.9), Carbon Monoxide (r=0.7), and Gas Eng.Emissions (r=0.6).

215. *Polycyclic Aromatic Hydrocarbons (Other).* Polycyclic aromatic hydrocarbons are a group of chemicals made up of three or more benzene rings interlinked in various arrangements. They are naturally present in fossil fuels or can be formed by thermal decomposition of any organic material containing carbon and hydrogen. This category was used to assign exposures to PAHs when the source material did not correspond to one of the three specific categories: coal, wood, petroleum. PAHs present in pyrolysis products of plastic, paint, rubber, food or other organic compounds would have been coded here. Main occupations: welders and flame cutters; roofers; chefs and cooks. Automatics: PAHs from any source. Lifetime prevalence: 19.4%. Top three positively correlated chemicals: Other Pyrolysis Fumes (r=0.9), Metal Oxide Fumes (r=0.6), and Iron Compounds (r=0.5).

216. *Polycyclic Aromatic Hydrocarbons (Wood).* Polycyclic aromatic hydrocarbons are a group of chemicals made up of three or more benzene rings interlinked in various arrangements. They are naturally present in fossil fuels or can be formed by thermal decomposition of any organic material containing carbon and hydrogen. Wood combustion is an important source of benzo(a)pyrene (which is coded separately) and of other PAHs. In fact, the quantity of benzo(a)pyrene in soots produced during wood combustion and coal combustion can be of the same order of magnitude. Main occupations: farmers; firefighters; chefs and cooks. Automatics: PAHs from any source. Lifetime prevalence: 4.2%. Top three positively correlated chemicals: PAH (Wood) (r=1.0), Phosgene (r=0.4), and Benzo(a)pyrene (r=0.3).

217. *Polycyclic Aromatic Hydrocarbons (Petroleum).* Polycyclic aromatic hydrocarbons are a group of chemicals made up of three or more benzene rings interlinked in various arrangements. They are naturally present in fossil fuels or can be formed by thermal decomposition of any organic material containing carbon and hydrogen. Crude oil, certain petroleum-derived substances (e.g., heavy fuel oil, asphalt, etc.) and their combustion products contain PAHs, albeit in smaller quantities than similar coal-derived products. Furthermore, concentrations of PAHs may increase in some of these products during use (e.g., used motor oils). Main occupations: Motor vehicle drivers and salesmen; motor vehicle repairmen; machinists. Automatics: PAHs from any source. Lifetime prevalence: 60.6%. Top three positively correlated chemicals: PAH (Any) (r=0.9), Gas Eng.Emissions (r=0.7), and Carbon Monoxide (r=0.6).

218. *Polycyclic Aromatic Hydrocarbons (Coal).* Polycyclic aromatic hydrocarbons are a group of chemicals made up of three or more benzene rings interlinked in various arrangements. They are naturally present in fossil fuels or can be formed by thermal decomposition of any organic material containing carbon and hydrogen. The most important sources of PAHs (in terms of quantity of PAHs released on a weight percent basis) in the workplace are coal tar products. Coal tar products are used at high temperatures in many industrial processes such as aluminium smelting and iron and steel production, thereby emitting PAHs into the work atmosphere. Main occupations: stationary engineers and boiler room workers; railway trackmen; pipefitters and plumbers. Automatics: PAHs from any source. Lifetime prevalence: 7.6%. Top three positively correlated chemicals: Coal Comb.Products (r=0.8), Coal Tar and Pitch (r=0.5), and Benzo(a)pyrene (r=0.5).

219. *Benzo(a)pyrene.* A number of five and six-membered ring PAHs are regarded as being carcinogenic and among these, benzo(a)pyrene, a six ring compound, has been the subject of special interest. It is the most frequently studied PAH and analytical methods for its determination have been available for a long time; in fact, determination of the benzo(a)pyrene

exposure has often been used as a proxy for PAH exposure. The highest concentrations of benzo(a)pyrene occur in coal tar products. Main occupations: motor vehicle mechanics; machinists; foundry workers. Automatics: PAHs from any source. Lifetime prevalence: 21.2%. Top three positively correlated chemicals: Soot (r=0.6), Sulphur Dioxide (r=0.5), and PAH (Coal) (r=0.5).

220. *Monocyclic Aromatic Hydrocarbons*. MAHs are those hydrocarbons that possess the special properties associated with the benzene nucleus or ring, in which six carbon-hydrogen groups are arranged at the corners of the hexagon. This includes all aromatic compounds that have only one benzene ring including substituted products such as xylene, toluene, styrene (all of which were also coded separately), phenol and ethyl benzene and others. These substances are present in certain petroleums, solvents, motor and heating fuels and coal tar distillates. Main occupations: motor vehicle mechanics; metal machinists; painters. Lifetime prevalence: 34.4%. Top three positively correlated chemicals: Alkanes (C5-C17) (r=0.8), Benzene (r=0.6), and Mineral Spirits+BTX (r=0.6).

221. *Aromatic Alcohols*. Includes all phenolic derivatives such as phenol itself (which was also coded separately), hydroquinone (a reducing agent in photographic developing baths), o-phenyl phenol (contained in Lysol®), hexachlorophene (Phisohex®, an antibacterial detergent used in hospitals), pyrocatechol (an antioxidant in rubber) and resorcinol (used in tanning). Creosote (also coded separately), obtained from coal tar, also contains a small amount of phenol. Main occupations: janitors; railway trackmen; roofers. Lifetime prevalence: 3.0%. Top three positively correlated chemicals: Phenol (r=0.7), Creosote (r=0.6), and Phenol-Formald. (r=0.3).

222. *Aromatic Amines*. Aromatic hydrocarbons in which at least one hydrogen atom has been substituted by a primary, secondary or tertiary amino group. The main compounds coded here are aniline, benzidine and naphtylamines. Other chemical groups may also be present on the aromatic ring. Many dyes and organic pigments contain the aromatic amine function. Main occupations: painters; printshop workers; shoemakers and repairmen. Lifetime prevalence: 6.7%. Top three positively correlated chemicals: Organic Dyes & Pig. (r=0.7), Carbon Black (r=0.5), and Inorg.Pigments (r=0.5).

223. *Phthalates*. Because of their low vapor pressures and chemical stability, various esters of phthalic acid such as diethyl, dibutyl and di-n-octyl are used as plasticizers to impart flexibility to certain plastics, notably polyvinyl chloride or PVC. Main occupations: motor vehicle refinishers; plastics workers; carpenters. Lifetime prevalence: 2.3%. Top three positively correlated chemicals: Cellulose Nitrate (r=0.5), Aliphatic Esters (r=0.4), and Titanium Dioxide (r=0.4).

224. *Isocyanates*. Represented by the general formula R-N=C=O, these are basic constituents in the production of polyurethanes, which in turn are used as flexible and rigid foams and in resins, paints and varnishes. Toluene diisocyanate (TDI) is the most commonly used of the isocyanates but all other types were included. Main occupations: motor vehicle refinishers; motor vehicle mechanics; foundry workers. Lifetime prevalence: 1.6%. Top three positively correlated chemicals: Styrene (r=0.6), Polyurethanes (r=0.5), and Phosgene (r=0.5).

225. *Cleaning Agents*. Materials which have cleansing action such as soap. Their main function is to aid water in the cleaning process. They may be simple sulphonated fatty acids or complex synthetic materials. Organic solvents were excluded here and have been coded separately. Main occupations: janitors; chefs and cooks; restaurant busboys. Lifetime prevalence: 16.1%. Top three positively correlated chemicals: Javel Water (r=0.4), Hypochlorites (r=0.4), and Biocides (r=0.4).

226. *Pharmaceuticals*. All products used as sedatives, tranquilizers, narcotics, painkillers, and all other prescription and non-prescription drugs and remedies. Veterinary medicines were also included. Main occupations: pharmacists; physicians and surgeons; pharmaceutical industry workers. Lifetime prevalence: 1.6%. Top three positively correlated chemicals: Laboratory Products (r=0.2), Biocides (r=0.1), and Mercury Compounds (r=0.1).

227. *Laboratory Products*. A general category used to code exposure to a laboratory environment. Main occupations: chemical engineers; chemists; physical science technicians. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Nitric Acid (r=0.2), Acetone (r=0.2), and Ethanol (r=0.2).

228. *Fertilizers*. Materials that are added to the soil to supply plant food either directly or by chemical reaction with the soil. Commercial fertilizers include nitrates, phosphates, potash salts, calcium salts or mixtures of these. Main occupations: farmers; nursery workers; dockworkers. Lifetime prevalence: 5.6%. Top three positively correlated chemicals: Pesticides (r=0.7), Grain Dust (r=0.4), and Arsenic Compounds (r=0.3).

229. *Pesticides*. Substances capable of killing some form of organism that is deemed to be undesirable. Pesticides include insecticides, herbicides, rodenticides, fungicides, molluscicides and nematodicides. Farming is the main occupation in which pesticides are used in large quantities, and are likely to be handled in an unsafe manner. Main occupations: farmers; dockworkers; nursery workers. Lifetime prevalence: 5.6%. Top three positively correlated chemicals: Fertilizers (r=0.7), Grain Dust (r=0.4), and Arsenic Compounds (r=0.4).

230. *Biocides*. Includes all products used to disinfect, deodorize, sterilize and sanitize. This implies the capability of killing micro-organisms (algae, bacteria, viruses, etc.). This group therefore includes bactericides, algicides, fungicides, germicides and preservatives. Agricultural pesticides were coded separately. Main occupations: janitors; painters; barbers and hairdressers. Lifetime prevalence: 9.5%. Top three positively correlated chemicals: Hypochlorites (r=0.5), Javel Water (r=0.5), and Ammonia (r=0.4).

231. *Bleaches*. Substances or mixtures which have the ability to chemically remove dyes or pigments that exist naturally in a material or that have been added to it in an industrial process. They are widely used in the treatment of cellulose, in the pulp and paper industry and of course in the textile industry. The main active agents found in bleaches include chlorine, calcium hypochlorite, potassium hypochlorite, sodium hypochlorite (Javex), chlorine dioxide, sodium chlorate, hydrogen peroxide and detergents (several of which have also been coded separately). Main occupations: laundry workers and dry cleaners; photographers; photographic processors. Lifetime prevalence: 1.1%. Top three positively correlated chemicals: Hypochlorites (r=0.4), Chlorine (r=0.4), and Javel Water (r=0.4).

3. SAS matrix language program for semi-Bayes modeling

The SAS-IML program below is taken directly from Witte et al. (1998) with errata corrected. Changes were made to incorporate the present data and prior, but the semi-Bayes estimation portions of the program were left unchanged. Whereas the original program also calculated empirical-Bayes estimates, the program below refers only to semi-Bayes estimation.

B is the n by 1 column vector of beta coefficients from a logistic regression, n being the number of determinants; Z is the n by k matrix of the prior information, k being the number of second-level covariates; V is the n by n matrix of beta covariance values; and T2 is the n by 1 vector of specified prior variances.

```
/* Initialize variables: */
np=nrow(Z);                    /* Number of first-stage parameters. */
ncol=ncol(Z);                  /* Number of second-stage parameters, including
                                  an intercept term. */
df2=np-ncol(Z);                /* Second-stage degrees of freedom. */
Inp=I(np);                     /* Creating an n by n identity matrix */
max_c=50; count=0;             /* Maximum number and count of iterations. */
/* Undertake second-stage linear regression: */
w=inv(v+t2#Inp);               /* Weight matrix. */
wv=w*v; wz=w*z;
ws=sum(w);
vs=inv(t(Z)*w*Z);              /* Invert 2nd-stage information. */
bs=vs*(t(wz)*b);               /* 2nd-stage coefficient estimates. */
e=b-(Z*bs);                    /* Residual from 2nd-stage estimates. */
rsst=t(e)*w*e;                 /* Total residual sums of squares. */
rms=np*rsst/(df2*ws);          /* Residual mean square. */
/* Calculate posterior expectations of 1st-stage parameters: */
wvc=wv;
hatw=Z*vs*t(wz);               /* Projection of b to prior mean. */
st2=sqrt(t2*t(t2));
hatp=wvc*hatw + w#st2 + (wv-wvc);      /* Projection of b to posterior mean. */
```

```
vp=v-t(wvc)*(Inp-hatw)*v;          /* Estimated posterior covariance. */
npa=trace(hatp);                   /* Effective model degrees of freedom. */
bp=hatp*b;                         /* Estimated posterior mean. */
/* Calculate variances for interval estimates: */
varvp=vecdiag(vp);                 /* Transforming diagonal into an n by 1 vector, */
                                   /* of posterior variances. */
do i=1 to np;
stderr[i]=sqrt(varvp[i]);          /* standard errors of the posterior estimates */
bpn[i]=bp[i];                      /* the posterior estimates */
end;
```

Example of matrices

The following are example matrices like the ones I used in my analyses with the SAS-IML program. Assuming an analysis of five substances, the following represents the 5x1 column vector of fitted maximum likelihood coefficients, extracted from SAS PROC LOGISTIC output, B=[-0.139, -0.206, 0.363, 1.738, -0.683]; the 5x5 matrix of beta covariances,

$$V= \begin{bmatrix} 0.0333 & 0.0008 & -0.0025 & -0.0062 & -0.0002 \\ \quad ... & & & & \\ -0.0001 & 0.0014 & 0.0047 & -0.0030 & 0.1671 \end{bmatrix},$$

also extracted from PROC LOGISTIC using the COVOUT option; the 5x3 matrix of prior information, here representing a continuous intercept term for previous evidence and two other categories of exchangeability coded dichotomously as either 0 or 1,

$$Z = \begin{bmatrix} 0.18 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0.64 & 0 & 1 \end{bmatrix};$$

and the 5x1 column vector of prior variances, all set to reflect a tenfold range, T2=[0.35, 0.35, 0.35, 0.35, 0.35].

4. Results from using different definitions of the control series

Two different options for defining the control series in the study were available from the Montreal study. Patients with cancer other than of the lung were used as the control series for the main body of results in the thesis. This section provides results using the alternative option, which was comprised of 533 men identified from either an electoral list or random digit dialling procedure. A third option is also included, involving an equal-weight combination of the cancer and electoral-list series: a random sample of 533 men with cancer diagnoses were added to the 533 men of the electoral list series. Table A-4 provides rate ratios and 90% confidence limits for all 231 chemicals under these different design options. All results correspond to modeling strategy 2, where the effect of each chemical was adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

Table A-4: Rate ratio estimates, from strategy 2, for 231 chemicals, at ANY level of exposure, using three different designs for the control series

		Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
1.	Abrasives Dust	237	1.2 (1.0, 1.4)	1.2 (0.9, 1.5)	1.2 (1.0, 1.5)
2.	Inorg.Insul.Dust	112	1.1 (0.9, 1.4)	1.1 (0.8, 1.5)	1.1 (0.8, 1.4)
3.	Excavation Dust	109	1.5 (1.2, 1.8)	1.8 (1.2, 2.7)	1.5 (1.1, 2.0)
4.	Metallic Dust	276	1.3 (1.1, 1.6)	1.4 (1.1, 1.8)	1.3 (1.1, 1.6)
5.	Asbestos	177	1.2 (1.0, 1.4)	1.3 (1.0, 1.7)	1.2 (1.0, 1.5)
6.	Crystalline Silica	238	1.3 (1.1, 1.5)	1.3 (1.0, 1.7)	1.3 (1.1, 1.6)
7.	Portland Cement	79	1.4 (1.0, 1.8)	1.6 (1.1, 2.5)	1.6 (1.2, 2.3)
8.	Glass Dust	18	2.0 (1.1, 3.5)	2.8 (1.0, 7.4)	2.1 (1.0, 4.3)
9.	Glass Fibres	50	0.9 (0.7, 1.2)	0.9 (0.6, 1.4)	0.9 (0.6, 1.3)
10.	Industrial Talc	35	0.9 (0.6, 1.2)	0.8 (0.5, 1.2)	0.9 (0.6, 1.4)
11.	Brick Dust	34	0.9 (0.6, 1.3)	1.8 (1.0, 3.4)	1.6 (0.9, 2.6)
12.	Clay Dust	28	1.9 (1.2, 3.0)	1.5 (0.8, 2.7)	2.0 (1.2, 3.5)
13.	Concrete Dust	97	1.2 (0.9, 1.5)	1.1 (0.7, 1.5)	1.1 (0.8, 1.5)
14.	Bronze Dust	11	1.0 (0.5, 1.9)	0.9 (0.4, 2.3)	1.1 (0.5, 2.3)
15.	Brass Dust	24	1.6 (1.0, 2.6)	1.6 (0.8, 3.0)	1.6 (0.9, 2.8)
16.	Stainless Steel Dust	51	1.6 (1.2, 2.2)	1.3 (0.8, 2.1)	1.1 (0.8, 1.6)
17.	Mild Steel Dust	169	1.3 (1.1, 1.6)	1.1 (0.8, 1.5)	1.2 (1.0, 1.5)
18.	Inorg.Pigments	93	1.3 (1.0, 1.6)	1.2 (0.9, 1.8)	1.1 (0.8, 1.4)
19.	Mineral Wool Fibres	61	1.1 (0.8, 1.5)	1.2 (0.8, 1.8)	1.1 (0.8, 1.6)
20.	Extenders	49	1.0 (0.7, 1.3)	1.0 (0.7, 1.6)	0.9 (0.6, 1.3)
21.	Aluminium Alloy Dust	63	1.5 (1.1, 2.0)	0.9 (0.6, 1.3)	0.9 (0.7, 1.3)

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
22. Ashes	36	1.4 (0.9, 2.1)	1.0 (0.6, 1.7)	1.2 (0.7, 1.9)
23. Cosmetic Talc	13	1.3 (0.7, 2.4)	1.1 (0.4, 2.5)	1.4 (0.7, 3.0)
24. Borates	11	2.0 (1.0, 3.9)		6.4 (2.0, 20.2)
25. Sodium Carbonate	15	1.2 (0.7, 2.1)	1.0 (0.5, 2.1)	1.1 (0.6, 2.0)
26. Alumina	160	1.3 (1.1, 1.6)	1.2 (0.9, 1.7)	1.2 (1.0, 1.6)
27. Silicon Carbide	51	1.1 (0.8, 1.5)	0.9 (0.6, 1.5)	0.8 (0.6, 1.2)
28. Sulfur	9	0.7 (0.4, 1.4)	0.5 (0.2, 1.1)	0.6 (0.3, 1.2)
29. Calcium Oxide	69	1.1 (0.8, 1.4)	0.9 (0.6, 1.4)	1.2 (0.9, 1.7)
30. Calcium Sulphate	100	1.2 (0.9, 1.5)	1.1 (0.8, 1.5)	1.0 (0.8, 1.4)
31. Calcium Carbonate	46	1.0 (0.7, 1.3)	0.7 (0.5, 1.1)	0.7 (0.5, 1.0)
32. Titanium Dioxide	38	1.1 (0.8, 1.6)	1.2 (0.7, 2.0)	1.0 (0.6, 1.5)
33. Iron Dust	37	1.1 (0.7, 1.6)	1.4 (0.9, 2.4)	1.1 (0.7, 1.6)
34. Iron Oxides	101	1.1 (0.9, 1.4)	1.2 (0.9, 1.7)	1.1 (0.9, 1.5)
35. Copper Dust	47	1.3 (0.9, 1.8)	1.5 (0.9, 2.4)	1.2 (0.8, 1.7)
36. Zinc Dust	26	1.6 (1.0, 2.6)	1.3 (0.7, 2.5)	1.4 (0.8, 2.4)
37. Zinc Oxide	34	1.2 (0.8, 1.8)	1.2 (0.7, 2.1)	1.0 (0.6, 1.5)
38. Lead Oxides	22	1.8 (1.1, 2.9)	1.6 (0.8, 3.2)	1.4 (0.8, 2.4)
39. Basic Lead Carb.	28	1.4 (0.9, 2.2)	1.1 (0.6, 2.0)	1.0 (0.6, 1.7)
40. Lead Chromate	35	1.1 (0.8, 1.7)	1.1 (0.7, 2.0)	1.1 (0.7, 1.7)
41. Organic Dyes & Pig.	70	1.0 (0.8, 1.3)	1.2 (0.8, 1.8)	1.0 (0.7, 1.4)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
42. Cotton Dust	65	0.9 (0.7, 1.2)	1.5 (0.9, 2.3)	1.0 (0.7, 1.4)
43. Wool Fibres	41	0.9 (0.6, 1.3)	1.3 (0.8, 2.2)	1.1 (0.7, 1.6)
44. Wood Dust	227	1.2 (1.0, 1.4)	0.9 (0.7, 1.1)	1.0 (0.8, 1.2)
45. Grain Dust	60	0.9 (0.7, 1.2)	0.9 (0.6, 1.3)	0.9 (0.6, 1.3)
46. Flour Dust	36	1.0 (0.7, 1.4)	1.0 (0.6, 1.6)	1.1 (0.7, 1.7)
47. Fur Dust	14	1.2 (0.7, 2.1)	1.1 (0.4, 2.5)	1.0 (0.5, 2.0)
48. Hair Dust	9	0.9 (0.5, 1.9)	0.9 (0.3, 2.7)	1.4 (0.6, 3.6)
49. Starch Dust	14	1.4 (0.8, 2.6)	0.9 (0.4, 1.9)	1.2 (0.6, 2.4)
50. Sugar Dust	15	1.6 (0.9, 3.0)	1.1 (0.5, 2.6)	1.5 (0.7, 3.0)
51. Leather Dust	21	0.7 (0.5, 1.1)	1.0 (0.5, 2.1)	0.7 (0.4, 1.2)
52. Tobacco Dust	9	1.0 (0.5, 2.1)	0.7 (0.3, 1.8)	1.0 (0.4, 2.2)
53. Natural Rubber	44	1.2 (0.8, 1.7)	1.4 (0.8, 2.4)	1.2 (0.8, 1.8)
54. Synthetic Fibres	45	0.9 (0.7, 1.2)	0.9 (0.5, 1.4)	1.0 (0.7, 1.4)
55. Plastic Dust	43	0.9 (0.7, 1.3)	1.3 (0.8, 2.1)	0.9 (0.6, 1.3)
56. Rayon Fibres	20	0.9 (0.6, 1.5)	1.3 (0.6, 2.8)	1.1 (0.6, 2.0)
57. Acrylic Fibres	18	0.8 (0.5, 1.4)	2.9 (0.9, 8.8)	1.4 (0.7, 2.8)
58. Polyester Fibres	31	1.0 (0.7, 1.5)	1.8 (0.9, 3.4)	1.2 (0.8, 2.0)
59. Nylon Fibres	23	1.1 (0.7, 1.7)	1.1 (0.5, 2.3)	1.1 (0.6, 2.0)
60. Acetate Fibres	12	0.8 (0.4, 1.5)	0.6 (0.3, 1.5)	0.8 (0.4, 1.6)
61. Cellulose Nitrate	18	0.7 (0.4, 1.2)	0.9 (0.5, 1.7)	0.7 (0.4, 1.2)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
62. Polyvinyl Chloride	11	0.6 (0.3, 1.2)	1.3 (0.5, 3.4)	0.5 (0.2, 0.9)
63. Polyvinyl Acetate	21	0.7 (0.4, 1.0)	0.8 (0.4, 1.5)	0.8 (0.5, 1.4)
64. Poly-Acrylates	29	1.3 (0.8, 1.9)	1.6 (0.9, 2.9)	1.4 (0.8, 2.2)
65. Alkyds	39	1.0 (0.7, 1.5)	1.6 (0.9, 2.8)	1.1 (0.7, 1.7)
66. Epoxies	13	1.8 (1.0, 3.5)	2.9 (1.0, 8.2)	2.4 (1.1, 5.5)
67. Phenol-Formald.	47	1.5 (1.0, 2.1)	0.8 (0.5, 1.3)	1.1 (0.7, 1.6)
68. Urea-Formald.	50	1.5 (1.1, 2.1)	0.9 (0.6, 1.5)	1.2 (0.8, 1.7)
69. Polyurethanes	18	1.4 (0.8, 2.4)	1.9 (0.7, 4.9)	1.7 (0.8, 3.5)
70. Styrene-Buta.Rubber	38	0.9 (0.6, 1.3)	0.9 (0.6, 1.5)	0.8 (0.6, 1.3)
71. Polychloroprene	30	1.1 (0.7, 1.6)	1.1 (0.6, 1.9)	1.2 (0.8, 1.9)
72. Fabric Dust	77	0.9 (0.7, 1.2)	0.9 (0.6, 1.3)	0.9 (0.6, 1.2)
73. Coal Dust	63	1.4 (1.0, 1.9)	1.0 (0.6, 1.4)	0.9 (0.7, 1.3)
74. Carbon Black	52	1.3 (0.9, 1.7)	1.1 (0.7, 1.7)	1.2 (0.8, 1.7)
75. Cellulose	58	1.1 (0.8, 1.5)	0.8 (0.6, 1.2)	1.0 (0.7, 1.4)
76. Soot	91	1.2 (0.9, 1.5)	1.1 (0.8, 1.6)	1.1 (0.8, 1.5)
77. Rubber Dust	31	0.9 (0.6, 1.3)	1.2 (0.7, 2.2)	1.0 (0.6, 1.5)
78. Graphite Dust	8	0.7 (0.3, 1.4)	1.4 (0.4, 4.5)	1.2 (0.5, 3.0)
79. Hydrogen	19	0.9 (0.5, 1.4)	1.6 (0.7, 3.7)	1.3 (0.7, 2.6)
80. Carbon Monoxide	478	1.2 (1.0, 1.4)	1.1 (0.9, 1.3)	1.1 (1.0, 1.4)
81. Hydrogen Cyanide	14	0.9 (0.5, 1.6)	2.2 (0.8, 5.8)	1.3 (0.6, 2.6)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
82. Ammonia	86	0.9 (0.7, 1.2)	1.3 (0.9, 1.9)	1.0 (0.8, 1.4)
83. Nitrogen Oxides	240	1.6 (1.3, 1.9)	1.5 (1.2, 1.9)	1.4 (1.1, 1.7)
84. Ozone	67	1.5 (1.1, 2.0)	1.6 (1.1, 2.5)	1.3 (1.0, 1.9)
85. Hydrogen Fluoride	38	2.0 (1.3, 2.9)	1.8 (1.0, 3.2)	1.9 (1.2, 3.1)
86. Sulphur Dioxide	144	1.1 (0.9, 1.3)	1.2 (0.9, 1.7)	1.1 (0.8, 1.3)
87. Hydrogen Sulphide	37	1.0 (0.7, 1.5)	0.9 (0.6, 1.6)	1.0 (0.6, 1.5)
88. Chlorine	15	0.5 (0.3, 0.9)	0.8 (0.4, 1.9)	0.6 (0.4, 1.2)
89. Hydrogen Chloride	59	1.0 (0.8, 1.3)	1.2 (0.8, 1.9)	1.0 (0.7, 1.4)
90. Natural Gas	24	1.0 (0.6, 1.5)	0.7 (0.4, 1.3)	0.8 (0.5, 1.3)
91. Methane	41	1.0 (0.7, 1.4)	0.7 (0.5, 1.2)	0.8 (0.5, 1.2)
92. Propane	39	1.2 (0.9, 1.8)	0.7 (0.5, 1.2)	0.9 (0.6, 1.4)
93. Formaldehyde	125	0.9 (0.7, 1.1)	1.0 (0.8, 1.4)	1.1 (0.8, 1.4)
94. Acetylene	47	1.6 (1.1, 2.2)	1.6 (1.0, 2.6)	1.6 (1.1, 2.4)
95. Phosgene	11	0.8 (0.4, 1.4)	1.5 (0.6, 4.0)	0.9 (0.4, 1.8)
96. Spray Gases	15	1.0 (0.6, 1.7)	0.9 (0.4, 1.8)	1.2 (0.6, 2.4)
97. Coal Gas	8	0.6 (0.3, 1.2)	0.7 (0.2, 2.1)	0.8 (0.3, 2.0)
98. Gas Welding Fumes	115	1.5 (1.2, 1.8)	1.9 (1.4, 2.7)	1.7 (1.3, 2.2)
99. Arc Welding Fumes	107	1.2 (0.9, 1.5)	1.1 (0.8, 1.5)	1.0 (0.8, 1.4)
100. Soldering Fumes	55	1.1 (0.8, 1.5)	1.4 (0.9, 2.2)	1.2 (0.8, 1.7)
101. Metal Oxide Fumes	190	1.3 (1.1, 1.5)	1.6 (1.2, 2.1)	1.4 (1.2, 1.8)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
102. Aluminium Fumes	23	1.4 (0.9, 2.3)	1.0 (0.5, 2.0)	0.9 (0.5, 1.5)
103. Calcium Oxide Fumes	66	1.3 (1.0, 1.8)	1.3 (0.9, 1.9)	1.2 (0.9, 1.7)
104. Chromium Fumes	43	2.2 (1.5, 3.3)	1.7 (1.0, 2.9)	1.6 (1.0, 2.4)
105. Manganese Fumes	60	1.6 (1.2, 2.2)	1.3 (0.8, 1.9)	1.2 (0.9, 1.7)
106. Iron Fumes	94	1.4 (1.1, 1.8)	1.4 (1.0, 1.9)	1.3 (1.0, 1.7)
107. Nickel Fumes	42	2.1 (1.4, 3.1)	2.0 (1.1, 3.4)	1.6 (1.1, 2.5)
108. Copper Fumes	47	2.2 (1.5, 3.1)	1.9 (1.1, 3.4)	1.8 (1.1, 2.7)
109. Zinc Fumes	39	1.6 (1.1, 2.3)	1.3 (0.8, 2.2)	1.6 (1.0, 2.5)
110. Silver Fumes	15	1.4 (0.8, 2.4)	2.1 (0.8, 5.6)	1.4 (0.7, 2.7)
111. Tin Fumes	49	1.6 (1.1, 2.2)	1.7 (1.0, 2.8)	1.6 (1.1, 2.4)
112. Lead Fumes	41	1.4 (1.0, 2.0)	1.4 (0.9, 2.3)	1.5 (1.0, 2.3)
113. Other Pyrolysis Fumes	171	1.3 (1.1, 1.6)	1.2 (0.9, 1.5)	1.2 (1.0, 1.5)
114. Cooking Fumes	57	0.8 (0.6, 1.1)	0.7 (0.4, 1.0)	0.8 (0.6, 1.1)
115. Gas Eng.Emissions	379	0.9 (0.8, 1.1)	0.8 (0.7, 1.0)	0.9 (0.8, 1.1)
116. Coal Comb.Products	51	1.4 (1.0, 2.0)	1.3 (0.8, 2.0)	1.3 (0.9, 1.9)
117. Diesel Eng.Emissions	165	1.2 (1.0, 1.5)	1.4 (1.1, 1.9)	1.3 (1.0, 1.7)
118. Liquid Fuel Comb.Prod.	71	1.2 (0.9, 1.6)	1.5 (1.0, 2.4)	1.6 (1.1, 2.2)
119. Wood Comb.Products	40	1.0 (0.7, 1.4)	1.1 (0.7, 2.0)	1.2 (0.8, 1.9)
120. Natural Gas Comb.Prod.	23	0.7 (0.5, 1.1)	0.7 (0.4, 1.3)	0.7 (0.4, 1.2)
121. Jet Fuel Eng.Emiss.	3	0.4 (0.2, 1.3)	0.7 (0.2, 3.5)	0.8 (0.2, 2.8)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
122. Propane Eng.Emiss.	28	1.7 (1.1, 2.6)	3.8 (1.6, 9.2)	2.7 (1.4, 5.2)
123. Plastics Pyrol.Prod.	17	0.6 (0.4, 0.9)	0.9 (0.5, 1.9)	0.6 (0.3, 1.0)
124. Rubber Pyrol.Prod.	20	0.9 (0.6, 1.5)	1.3 (0.6, 2.9)	0.9 (0.5, 1.6)
125. Propane Comb.Prod.	30	1.1 (0.7, 1.6)	0.6 (0.4, 1.0)	0.8 (0.5, 1.2)
126. Inorg.Acid Solutions	129	1.2 (1.0, 1.5)	1.3 (1.0, 1.8)	1.2 (0.9, 1.5)
127. Alkali, Caustic Solutions	72	1.3 (1.0, 1.7)	1.3 (0.9, 1.9)	1.2 (0.9, 1.7)
128. Javel Water	44	0.7 (0.5, 1.0)	0.6 (0.4, 1.0)	0.7 (0.5, 1.0)
129. Plating Solutions	10	1.1 (0.5, 2.2)	1.3 (0.5, 3.9)	1.3 (0.5, 2.9)
130. Nitric Acid	9	0.8 (0.4, 1.6)	3.2 (0.9, 11.3)	1.3 (0.6, 3.1)
131. Phosphoric Acid	14	1.7 (0.9, 3.1)	3.2 (1.1, 9.2)	1.9 (0.8, 4.2)
132. Sulphuric Acid	90	1.0 (0.8, 1.3)	1.4 (0.9, 2.0)	1.0 (0.8, 1.4)
133. Methanol	44	1.0 (0.7, 1.3)	0.7 (0.4, 1.1)	0.7 (0.5, 1.0)
134. Ethanol	15	1.4 (0.8, 2.5)	0.7 (0.4, 1.5)	1.0 (0.5, 1.9)
135. Ethylene Glycol	38	0.9 (0.6, 1.2)	0.8 (0.5, 1.3)	0.8 (0.5, 1.2)
136. Isopropanol	40	1.1 (0.8, 1.6)	1.0 (0.6, 1.6)	1.0 (0.6, 1.5)
137. Acetic Acid	28	0.9 (0.6, 1.4)	0.9 (0.5, 1.5)	0.8 (0.5, 1.3)
138. Carbon Tetrachloride	36	1.0 (0.7, 1.4)	0.9 (0.5, 1.4)	1.0 (0.6, 1.5)
139. Methylene Chloride	17	0.9 (0.6, 1.6)	0.8 (0.4, 1.6)	1.0 (0.5, 1.8)
140. 1,1,1.-Trichlorethane	16	1.8 (1.0, 3.3)	1.3 (0.6, 3.2)	1.3 (0.6, 2.5)
141. Trichloroethylene	25	1.3 (0.8, 2.0)	1.5 (0.7, 3.0)	1.5 (0.8, 2.6)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
142. Perchloroethylene	11	1.1 (0.6, 2.1)	3.1 (1.0, 9.5)	2.6 (1.1, 6.3)
143. Acetone	20	1.0 (0.6, 1.7)	0.9 (0.5, 1.6)	0.8 (0.5, 1.4)
144. Benzene	162	1.0 (0.8, 1.2)	1.2 (0.9, 1.6)	1.2 (1.0, 1.5)
145. Toluene	120	0.9 (0.8, 1.2)	1.2 (0.8, 1.6)	1.1 (0.8, 1.4)
146. Xylene	96	0.9 (0.8, 1.2)	1.1 (0.8, 1.5)	1.1 (0.8, 1.4)
147. Styrene	10	0.5 (0.3, 0.9)	0.8 (0.4, 2.0)	0.9 (0.4, 1.9)
148. Phenol	11	0.8 (0.4, 1.6)	0.9 (0.4, 2.1)	0.7 (0.4, 1.5)
149. Animal & Vege.Glues	34	1.0 (0.7, 1.4)	1.6 (0.9, 2.9)	1.2 (0.7, 1.8)
150. Turpentine	58	1.2 (0.9, 1.6)	1.0 (0.7, 1.6)	0.9 (0.7, 1.3)
151. Linseed Oil	53	1.3 (1.0, 1.8)	1.1 (0.7, 1.8)	0.9 (0.7, 1.4)
152. Synthetic Adhesives	133	1.0 (0.8, 1.2)	0.9 (0.7, 1.3)	1.0 (0.8, 1.2)
153. Solvents	375	1.2 (1.0, 1.4)	1.2 (1.0, 1.5)	1.2 (1.0, 1.4)
154. Waxes, Polishes	56	0.9 (0.7, 1.3)	1.1 (0.7, 1.8)	0.9 (0.6, 1.3)
155. Leaded Gasoline	122	1.1 (0.9, 1.4)	1.0 (0.7, 1.4)	1.2 (0.9, 1.5)
156. Kerosene	69	1.6 (1.2, 2.2)	1.1 (0.7, 1.7)	1.5 (1.0, 2.2)
157. Diesel Oil	45	1.4 (1.0, 1.9)	2.0 (1.1, 3.6)	1.6 (1.0, 2.4)
158. Heating Oil	53	1.4 (1.0, 2.0)	1.5 (0.9, 2.5)	1.6 (1.1, 2.4)
159. Mineral Spirits	110	1.2 (1.0, 1.5)	1.1 (0.8, 1.5)	1.1 (0.9, 1.4)
160. Lubric.Oils & Greases	291	1.2 (1.0, 1.4)	1.0 (0.8, 1.2)	1.0 (0.8, 1.2)
161. Cutting Fluids	85	1.3 (1.0, 1.7)	0.9 (0.6, 1.2)	0.9 (0.7, 1.2)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
162. Asphalt	30	0.9 (0.6, 1.3)	0.8 (0.4, 1.4)	0.7 (0.5, 1.1)
163. Coal Tar and Pitch	23	1.0 (0.7, 1.7)	0.9 (0.5, 1.8)	1.0 (0.6, 1.7)
164. Creosote	5	0.7 (0.3, 1.7)	0.3 (0.1, 0.9)	0.6 (0.2, 1.5)
165. Hydraulic Fluid	37	1.0 (0.7, 1.4)	0.9 (0.6, 1.5)	0.9 (0.6, 1.3)
166. Other Mineral Oils	32	1.0 (0.7, 1.4)	1.7 (0.9, 3.0)	1.3 (0.8, 2.2)
167. Jet Fuel	6	0.6 (0.3, 1.4)	0.5 (0.2, 1.5)	0.6 (0.2, 1.6)
168. Aviation Gasoline	6	0.6 (0.3, 1.2)	1.0 (0.3, 3.5)	0.7 (0.3, 1.6)
169. Mineral Spirits+BTX	158	1.3 (1.0, 1.5)	1.1 (0.8, 1.5)	1.2 (1.0, 1.5)
170. Cutting Fluids pre 1955	65	1.3 (1.0, 1.7)	1.0 (0.6, 1.4)	0.9 (0.7, 1.3)
171. Cutting Fluids post 1955	62	1.7 (1.2, 2.3)	1.0 (0.7, 1.5)	1.1 (0.8, 1.5)
172. Other Paints,Varnishes	124	1.1 (0.9, 1.4)	1.1 (0.8, 1.6)	1.0 (0.8, 1.3)
173. Wood Varnishes, Stains	56	1.2 (0.9, 1.7)	1.3 (0.8, 2.0)	1.3 (0.9, 1.8)
174. Inks	37	1.5 (1.0, 2.2)	1.1 (0.7, 1.9)	1.3 (0.8, 2.0)
175. Metal Coatings	74	1.2 (0.9, 1.6)	1.4 (0.9, 2.0)	1.4 (1.0, 2.0)
176. Cyanides	17	1.0 (0.6, 1.7)	2.1 (0.9, 5.1)	1.5 (0.8, 2.8)
177. Fluorides	42	1.8 (1.3, 2.7)	2.0 (1.1, 3.4)	1.9 (1.2, 3.0)
178. Chromium (VI) Comp.	90	1.4 (1.1, 1.8)	1.3 (0.9, 1.9)	1.3 (1.0, 1.7)
179. Hypochlorites	45	0.7 (0.5, 1.0)	0.6 (0.4, 1.0)	0.7 (0.5, 1.0)
180. Nitrates	8	0.9 (0.4, 1.9)	1.6 (0.5, 4.9)	1.3 (0.5, 3.2)
181. Beryllium Compounds	5	1.1 (0.4, 3.0)	0.8 (0.2, 2.8)	0.8 (0.3, 2.4)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
182. Magnesium Compounds	19	1.9 (1.1, 3.4)	1.2 (0.6, 2.5)	1.2 (0.7, 2.3)
183. Aluminium Compounds	199	1.4 (1.2, 1.7)	1.3 (1.0, 1.7)	1.3 (1.0, 1.6)
184. Titanium Compounds	44	1.1 (0.8, 1.6)	1.3 (0.8, 2.1)	1.0 (0.7, 1.5)
185. Vanadium Compounds	16	1.5 (0.8, 2.8)	0.9 (0.4, 2.0)	1.2 (0.6, 2.4)
186. Chromium Compounds	130	1.4 (1.1, 1.8)	1.3 (1.0, 1.8)	1.3 (1.0, 1.6)
187. Manganese Compounds	71	1.5 (1.1, 2.0)	1.1 (0.8, 1.6)	1.1 (0.8, 1.6)
188. Iron Compounds	248	1.2 (1.0, 1.5)	1.3 (1.0, 1.7)	1.2 (1.0, 1.5)
189. Cobalt Compounds	20	1.4 (0.9, 2.3)	0.7 (0.4, 1.2)	0.9 (0.6, 1.6)
190. Nickel Compounds	79	1.7 (1.3, 2.2)	1.6 (1.1, 2.3)	1.3 (0.9, 1.7)
191. Copper Compounds	128	1.3 (1.1, 1.6)	1.4 (1.0, 1.9)	1.3 (1.0, 1.6)
192. Zinc Compounds	107	1.4 (1.1, 1.7)	1.2 (0.9, 1.7)	1.3 (1.0, 1.6)
193. Arsenic Compounds	31	0.8 (0.6, 1.2)	1.0 (0.5, 1.7)	0.9 (0.5, 1.4)
194. Silver Compounds	24	1.2 (0.8, 1.9)	1.4 (0.7, 2.7)	1.2 (0.7, 2.1)
195. Cadmium Compounds	11	1.7 (0.9, 3.5)	1.4 (0.5, 3.6)	1.5 (0.7, 3.4)
196. Tin Compounds	92	1.4 (1.1, 1.8)	1.6 (1.1, 2.3)	1.7 (1.2, 2.3)
197. Antimony Compounds	19	1.3 (0.8, 2.2)	1.9 (0.9, 4.0)	2.0 (1.1, 3.8)
198. Tungsten Compounds	11	1.4 (0.7, 2.6)	0.5 (0.3, 1.1)	0.7 (0.4, 1.4)
199. Gold Compounds	12	2.0 (1.0, 4.0)	1.2 (0.5, 2.9)	1.7 (0.7, 3.6)
200. Mercury Compounds	17	1.3 (0.7, 2.2)	1.4 (0.6, 3.0)	1.1 (0.6, 2.0)
201. Lead Compounds	434	1.0 (0.9, 1.2)	1.0 (0.8, 1.2)	1.0 (0.8, 1.2)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
202. Alkanes (C18+)	320	1.2 (1.0, 1.4)	1.0 (0.8, 1.3)	1.0 (0.9, 1.2)
203. Alkanes (C1-C4)	82	1.1 (0.8, 1.4)	0.7 (0.5, 1.1)	0.9 (0.7, 1.2)
204. Alkanes (C5-C17)	368	1.4 (1.2, 1.7)	1.4 (1.1, 1.8)	1.5 (1.2, 1.8)
205. Aliphatic Alcohols	83	1.0 (0.8, 1.3)	0.8 (0.6, 1.1)	0.9 (0.6, 1.1)
206. Aliphatic Aldehydes	146	0.8 (0.7, 1.0)	1.0 (0.8, 1.4)	1.0 (0.8, 1.3)
207. Chlorinated Alkanes	93	1.1 (0.9, 1.4)	0.9 (0.7, 1.3)	1.0 (0.7, 1.3)
208. Unsat.Aliph.Hydrocarb.	54	1.5 (1.1, 2.0)	1.5 (1.0, 2.4)	1.5 (1.0, 2.1)
209. Chlorinated Alkenes	42	1.2 (0.8, 1.7)	1.5 (0.9, 2.6)	1.4 (0.9, 2.2)
210. Aliphatic Esters	29	1.1 (0.7, 1.6)	1.3 (0.8, 2.2)	1.2 (0.7, 1.8)
211. Aliphatic Ketones	47	0.8 (0.6, 1.1)	0.9 (0.6, 1.3)	0.8 (0.5, 1.1)
212. Fluorocarbons	16	0.6 (0.4, 1.1)	0.6 (0.3, 1.1)	0.7 (0.4, 1.3)
213. Glycol Ethers	26	1.1 (0.7, 1.7)	1.3 (0.7, 2.5)	1.0 (0.6, 1.7)
214. PAH (Any)	581	1.1 (1.0, 1.4)	0.9 (0.7, 1.2)	1.0 (0.8, 1.2)
215. PAH (Other)	187	1.2 (1.0, 1.4)	1.1 (0.9, 1.4)	1.1 (0.9, 1.3)
216. PAH (Wood)	40	1.0 (0.7, 1.4)	1.1 (0.6, 1.9)	1.2 (0.8, 1.9)
217. PAH (Petroleum)	561	1.1 (1.0, 1.3)	0.9 (0.7, 1.2)	1.0 (0.8, 1.2)
218. PAH (Coal)	84	1.4 (1.0, 1.8)	1.1 (0.7, 1.5)	1.2 (0.9, 1.7)
219. Benzo(a)pyrene	220	1.2 (1.0, 1.4)	1.2 (0.9, 1.5)	1.1 (0.9, 1.4)
220. MAH	331	1.2 (1.0, 1.4)	1.2 (1.0, 1.5)	1.2 (1.0, 1.4)
221. Aromatic Alcohols	21	0.9 (0.6, 1.4)	0.5 (0.3, 0.9)	0.6 (0.4, 1.1)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.

	Exp Cases	Cancer Series (N=2172) RR (90% CL)	Electoral list series (N=533) RR (90% CL)	Combined series (N=1066) RR (90% CL)
222. Aromatic Amines	55	0.9 (0.7, 1.2)	1.4 (0.9, 2.2)	1.0 (0.7, 1.5)
223. Phthalates	15	0.5 (0.3, 0.9)	0.8 (0.4, 1.6)	0.7 (0.4, 1.2)
224. Isocyanates	16	1.0 (0.6, 1.7)	1.6 (0.6, 3.8)	1.3 (0.7, 2.6)
225. Cleaning Agents	154	1.0 (0.8, 1.2)	0.8 (0.6, 1.0)	0.9 (0.7, 1.1)
226. Pharmaceuticals	13	1.2 (0.7, 2.2)	1.2 (0.5, 2.6)	1.3 (0.7, 2.7)
227. Laboratory Products	7	1.1 (0.5, 2.2)	1.5 (0.5, 4.3)	1.8 (0.7, 4.8)
228. Fertilizers	57	1.2 (0.9, 1.6)	1.5 (0.9, 2.4)	1.3 (0.9, 1.9)
229. Pesticides	54	1.0 (0.7, 1.3)	0.9 (0.6, 1.4)	0.9 (0.6, 1.3)
230. Biocides	78	0.8 (0.6, 1.0)	0.8 (0.6, 1.1)	0.7 (0.6, 1.0)
231. Bleaches	6	0.5 (0.2, 1.1)	0.8 (0.3, 2.7)	0.6 (0.2, 1.5)

All estimates adjusted for eight non-occupational confounders: age, ethnicity, income, education, recreational activity, history of cigarette smoking, history of alcohol consumption, and respondent status.