Université de Montréal

Amélioration de l'évaluation de l'exposition professionnelle rétrospective dans les études épidémiologiques à base populationnelle

Par

Jean-François Sauvé

Département de santé environnementale et santé au travail

École de santé publique

Thèse présentée à l'École de santé publique

en vue de l'obtention du grade de docteur

en santé publique

de l'option Toxicologie et analyse du risque

Avril 2017

Résumé

L'évaluation de l'exposition professionnelle rétrospective forme une composante critique et complexe des études cas-témoins à base populationnelle, consistant à caractériser sans biais l'exposition de chaque sujet durant la carrière à partir d'informations limitées. Une approche développée à Montréal permettant d'estimer l'exposition à près de 300 agents par expertise à partir de descriptions d'emploi détaillées fut appliquée dans quatre études cas-témoins de cancer entre 1979-2005. La banque de données d'estimations accumulées par le groupe montréalais représente une source d'information unique au monde sur l'exposition passée pouvant servir au développement d'outils applicables à de nouvelles études. Cette recherche visait à développer et à caractériser des outils permettant d'améliorer l'évaluation de l'exposition professionnelle à travers l'exploitation de cette banque de données.

Le premier volet portait sur l'élaboration de la matrice emplois-expositions (MEE) CANJEM synthétisant l'information sur l'exposition à 258 agents issue de 31 673 emplois entre 1930-2005. CANJEM a été définie par 3 axes : agents, professions ou industries (selon 7 systèmes de classification possibles) et période temporelle (1, 2 ou 4 périodes). Chaque cellule (combination d'agent, de profession/industrie et de période) décrit le profil d'exposition des emplois au moyen de cinq indices : probabilité (pourcentage), fiabilité, intensité, et fréquence (en distributions relatives de catégories ordinales) et intensité d'exposition pondérée par la fréquence (continue). Plus de 90% de la population active canadienne est représentée dans les cellules de CANJEM selon 2 recensements (1986 et 2011).

Le deuxième volet visait à raffiner les estimations de CANJEM pour 5 agents par des modèles hiérarchiques bayésiens pour intégrer l'information sur l'exposition entre les cellules de

professions similaires, et en tenant compte de l'étude source des emplois. L'influence des autres professions sur les estimations des cellules comptant 1 à 4 emplois était parfois importante, tandis qu'elle était modérée ou négligeable à partir de 5 emplois. Le troisième volet visait à estimer des niveaux d'intensité quantitatifs de l'exposition aux poussières de bois aux cellules de CANJEM, en modélisant 5170 mesures historiques (1981-2003) de la banque *Canadian Workplace Exposure Database*. Les moyennes géométriques sur 8 heures prédites pour 1989 variaient entre 0,49 à 1,67 mg/m^3, avec des ratios de 1-1,3-2,3 entre les catégories faible, moyenne et élevée de CANJEM. Des niveaux quantitatifs ont pu être estimés pour toute profession avec une probabilité d'exposition non nulle dans CANJEM. Le dernier volet visait à comparer les expositions assignées par une approche hybride combinant expertise individuelle et profils d'emplois prédéfinis avec l'approche par expertise traditionnelle. Les comparaisons portant sur l'exposition à 203 agents pour 90 professions ont montré une augmentation de la fiabilité des évaluations telle que jugée par les experts. Une réduction dans la variabilité intra-profession des estimations a aussi été observée, potentiellement expliquée par une meilleure cohérence dans l'évaluation et par l'application des profils prédéfinis par profession.

L'exploitation de la banque d'expertises montréalaise a permis de développer une MEE multidimensionnelle et de la bonifier, notamment par des estimations quantitatives. CANJEM est une ressource unique pouvant être appliquée à l'évaluation et à la prévention des maladies professionnelles au Canada et ailleurs. L'évaluation de l'approche par expertise hybride a par ailleurs montré l'utilité d'exploiter les données existantes pour faciliter l'évaluation de l'exposition. En somme, cette thèse a permis d'élaborer un ensemble de méthodes transférables à l'exploitation d'autres banques de données d'évaluations rétrospectives.

Mots-clés : Exposition professionnelle, évaluation de l’exposition, matrices emplois-expositions, études cas-témoins, études populationnelles, études rétrospectives

Abstract

Retrospective occupational exposure assessment forms a critical and challenging component of population-based case-control studies that requires estimating lifetime exposures for each subject with high accuracy and limited information. In the 1980s, researchers in Montreal developed a novel method involving the expert review of detailed job histories to evaluate exposure to some 300 agents, which was applied to 4 large case-control studies of cancer in the Montreal region between 1979 and 2005. The data collected by the group represents a unique body of knowledge on retrospective occupational exposures that could inform the exposure assessment of new studies. The objective of this thesis was to develop and examine approaches designed to improve retrospective exposure assessment in population studies by building on the information contained in the Montreal group database.

The first component involved the development of CANJEM, a job-exposure matrix (JEM) summarizing the exposure data for 258 agents across 31,673 jobs spanning 1930-2005. CANJEM featured 3 axes: agents, occupation or industry (in 7 classification schemes, each in multiple resolutions) and periods (1, 2 or 4 categories). Each cell depicts the exposure profile of jobs using 5 indices: probability of exposure (percentage), reliability, intensity and frequency (as relative distributions of ordinal ratings) and frequency-weighted intensity (FWI; continuous). Over 90% of the Canadian working population were covered by CANJEM cells according to 2 national surveys (1986 and 2011).

The second component aimed at refining the CANJEM estimates for 5 agents by applying Bayesian hierarchical models to pool information on exposure among cells from similar occupations, while also accounting for the source studies of jobs. The estimates of cells based

on fewer than 5 jobs were often overly sensitive to the influence of the other occupations, whereas their influence ranged from moderate to negligible for cells based on 5 or more jobs. The third comportment aimed to develop quantitative estimates of wood dust exposure for CANJEM cells by modelling 5170 historical (1981-2003) measurements from the Canadian Workplace Exposure Database. Predicted geometric mean (GM) concentrations of cells for 8 hours, breathing zone and year 1989 ranged 0.49–1.67 mg/m^3, with contrasts of 1-1.3-2.3 in the GMs between the low, medium and high intensity ratings. The model provided estimates of wood dust concentrations for any cell with some exposure in CANJEM. The last component aimed to compare the exposures assigned with a novel hybrid approach combining expert assessment and job-exposure profiles summarizing past evaluations by occupation, to the traditional expert method. The comparisons of the exposure data covering 203 agents and 90 occupations showed a higher reliability in the assessments as rated by the experts with the hybrid method. A decrease in the within-occupation variability of the exposures assigned with the hybrid method was also found, which may be explained by a greater consistency in the assessments and by the application of the job-exposure profiles.

The use of the Montreal group database provided a backbone for developing and refining a multidimensional JEM, such as with the development of quantitative estimates. CANJEM represents a unique resource that could be applied to assess and prevent occupational diseases in Canada and elsewhere. The assessment of the hybrid expert approach also documented the usefulness of available exposure data to inform the process of exposure assessment. The set of methods developed in this project could also be applied to other retrospective exposure assessment databases.

Keywords: Occupational exposure, exposure assessment, job-exposure matrix, case-control studies, population studies, retrospective studies

Table des matières

Liste des tableaux

Chapitre 1

Chapitre 2

Chapitre 3

Chapitre 4

Chapitre 5

Chapitre 6

Chapitre 7

Annexe 1

Liste des figures

Chapitre 1

Chapitre 2

Chapitre 3

Chapitre 4

Chapitre 5

Chapitre 6

Liste des sigles et abréviations

AHS	Agricultural Health Study
BC	British Columbia
BMI	Body-mass index
CANJEM	Canadian job-exposure matrix
CCDO	Canadian Classification and Dictionary of Occupations
CCDP	Classification Canadienne Descriptive des Professions
CEHD	Chemical Exposure Health Data
CI	Confidence/credible interval
CIRC	Centre international de recherche sur le cancer
CITI	Classification internationale type, par industrie, de toutes les branches d'activité économique
CITP	Classification internationale type des professions
CNP	Classification nationale des professions
CTI	Classification-type des industries
CWED	Canadian Workplace Exposure Database
DRE	Digital rectal examination
FINJEM	Finnish job-exposure matrix
FWI	Frequency-weighted intensity
GM	Geometric mean
HAP	Hydrocarbures aromatiques polycycliques
IARC	International Agency for Research on Cancer
IMIS	Integrated management information system
IRSST	Institut de recherche Robert Sauvé en santé et en sécurité du travail

ISCO	International Standard Classification of Occupations
JAGS	Just Another Gibbs Sampler
JEM	Job-exposure matrix
JEP	Job-exposure profile
LOD	Limit of detection
MCMC	Markov-chain Monte Carlo
MEE	Matrice emplois-expositions
n.c.a.	Non classé autrement
n.e.c.	Not elsewhere classified
NAICS	North American Industry Classification System
NOC	National Occupational Classification
OR	Odds ratio
OSHA	Occupation Safety and Health Administration (États-Unis)
PAH	Polycyclic aromatic hydrocarbons
PCa	Prostate cancer
PSA	Prostatic specific antigen
RIE	Relative index of exposure
SB	Semi-Bayes
SCIAN	Système de classification des industries de l'Amérique du Nord
SIC	Standardized Industrial Classification
SOC	United States Standard Occupation Classification

Remerciements

J’aimerais d’abord remercier Jérôme Lavoué et Marie-Élise Parent pour leur confiance et de m’avoir permis de réaliser ce projet méthodologique un peu fou. Je les remercie infiniment pour avoir consacré leur expertise et leur temps précieux à travers toutes les étapes de ce travail, et pour m’avoir constamment poussé à regarder au-delà des aspects purement méthodologiques vers les objectifs fondamentaux d’identification et de prévention des maladies d’origines professionnelles.

Ce travail n’aurait jamais pu voir le jour sans le travail minutieux et la curiosité des experts-chimistes impliqués dans le développement et l’application de l’approche d’évaluation de l’exposition dans les études montréalaises. Je voudrais saluer particulièrement Louise Nadon, Ramzan Lakhani, Dora Rodriguez, Robert Bourbonnais et Mounia Rhazi, sans oublier Denis Bégin et Michel Gérin. Ce fut un plaisir de vous côtoyer et de connaître le côté humain et l’expertise qui se cache derrière les montagnes de données dépersonnalisées.

Je tiens à souligner l’apport fondamental de Jack Siemiatycki dans l’origine des études ayant servi de fondation à cette thèse et dans le développement de CANJEM. Je voudrais par ailleurs saluer les collègues du département de santé environnementale de l’Université de Montréal, du Centre de recherche du CHUM et du groupe d’épidémiologie et de biostatistique de l’Institut Armand-Frappier, qui ont soit contribué à la réalisation de cette thèse ou soit à me faire décrocher de celle-ci sur l’heure du midi ou lors d’autres occasions. Je remercie particulièrement Marie-Pierre Sylvestre pour m’avoir aiguillé vers de nouvelles approches statistiques qui, bien qu’elles n’aient pas toujours été fructueuses, m’ont permis d’élargir mon bagage de connaissances méthodologiques; Hughes Richard, pour les discussions éclairantes et pour son travail de moine dans la préparation des données; Dan Vatnik, grâce à qui j'ai pu apprivoiser les méthodes de calcul haute performance; et finalement Philipe Sarazin, avec qui j'ai pu partager les défis et les joies du cheminement doctoral.

Je voudrais également souligner l'appui des Fonds de Recherche du Québec-Santé, de l'Institut de recherche Robert-Sauvé en santé et en sécurité du travail, et de Faculté des études supérieures et postdoctorales de l'Université de Montréal, qui m'ont permis de me consacrer entièrement à la réalisation de ce travail au cours de ces cinq années.

J'aimerais aussi remercier mes parents, qui m'ont toujours encouragé dans mes études, et à qui je dois cette curiosité et cette débrouillardise essentielles à la recherche scientifique.

Finalement, je voudrais remercier Edith de tout mon cœur pour son soutien continuel et infaillible tout au long de mes études et dans les périodes de doute de remise en question. Son support m'a permis de sortir relativement indemne de ce long sprint final pour concrétiser cette thèse.

Chapitre 1. Mise en contexte

1.1 Introduction générale

La prévention des maladies d'origines environnementale et professionnelle repose d'une part sur la connaissance des impacts sur la santé reliée aux agents présents l'environnement et, d'autre part, sur la distribution des niveaux d'exposition à ces agents dans la population. L'analyse du risque toxicologique est un processus systématique qui permet de caractériser les risques à la santé reliés à l'exposition à un agent dans la population, et de mettre en place des normes pour maintenir l'exposition à des niveaux acceptables afin de protéger la population générale ou des groupes spécifiques tels les travailleurs.

La démarche d'évaluation de l'exposition dans la population, que ce soit par des mesures dans l'environnement ou dans les matrices biologiques, ou par des prédictions de scénarios d'exposition, entre dans ces deux dimensions puisqu'elle joue un rôle important dans la caractérisation du potentiel dangereux des agents environnementaux. L'identification du danger, en particulier quant à la cancérogénicité ou non d'une substance, repose sur une approche de valeur probante de la preuve (*weight of evidence*), la plus forte étant les associations observées par des études chez l'humain (National Research Council, 1983; EPA, 2005; Centre international de recherche sur le cancer, 2006). L'évaluation de l'exposition constitue une composante critique de ces études, soit la caractérisation de l'exposition des individus sur l'ensemble de leur vie ou de leur carrière, en raison du long temps de latence associé au développement de certaines maladies chroniques, tel que le cancer (Nieuwenhuijsen, 2003). L'évaluation de l'exposition passée peut permettre l'identification et la caractérisation de la relation entre l'exposition et la maladie, soit la relation dose-réponse (ou exposition-réponse). Celle-ci représente une valeur probante cruciale dans la détermination du potentiel cancérogène d'un agent (Steenland et Deddens, 2004), et facilite également la surveillance du fardeau des maladies dans la population et leur prévention, notamment par la mise en place de valeurs sanitaires.

Le caractère cancérogène d'un grand nombre de substances, par exemple l'amiante ou le benzène, a été découvert par l'entremise d'études chez les travailleurs. L'environnement professionnel représente encore aujourd'hui une fenêtre privilégiée pour caractériser le potentiel cancérogène de l'exposition à divers agents, dont plusieurs se retrouvent dans l'environnement général (Siemiatycki et coll., 2004). L'intérêt du milieu de travail est multiple : les niveaux d'exposition y sont communément plus élevés comparativement à ceux dans la population générale, bien que la durée de l'exposition peut être plus courte (Semple, 2005), et le potentiel d'information disponible pour caractériser l'exposition passée (p. ex. mesures d'exposition historiques, procédés industriels) peut y être plus étoffé (Nieuwenhuijsen et coll., 2006). L'évaluation des expositions subies par les sujets d'une étude au cours de leur carrière représente néanmoins un processus complexe, puisqu'elle doit concilier un objectif de justesse et de précision dans les estimations avec une base d'information de qualité et de quantité hétérogènes, qui varie également en fonction du devis des études.

1.1.1 Principaux devis utilisés en épidémiologie professionnelle

La découverte du potentiel cancérogène d'agents environnementaux a historiquement reposé sur l'identification d'une série (ou noyau) de cas de maladies rares chez des groupes de travailleurs. Notons par exemple le cancer du scrotum chez les ramoneurs de cheminées, rapporté par Percival Pott à la fin du 18e siècle (Pleil et coll., 2012), le cancer de la vessie chez les travailleurs de l'industrie des colorants (Case et coll., 1954), le cancer du nez associé à l'exposition à des aérosols de nickel (Siemiatycki et coll., 2006), et l'angiosarcome du foie chez les travailleurs exposés au chlorure de vinyle (Vineis et Blair, 1992; Kielhorn et coll., 2000). La faible taille d'échantillon et l'absence de groupe de comparaison des séries de cas représentent des limites majeures qui requièrent des devis plus élaborés permettant d'approfondir les associations observées (Blair et coll., 1996). De nouvelles hypothèses continuent néanmoins à être générées par cette approche, telles l'association entre l'exposition au diacétyle et l'apparition de bronchiolite constrictive (Parmet et Von Essen, 2002), ou le cholangiocarcinome chez des travailleurs de l'imprimerie au Japon (Kumagai et coll., 2013).

Les études de cohortes industrielles sont basées sur le suivi d'une population de travailleurs œuvrant dans une industrie ou une entreprise commune et partageant un profil d'exposition relativement homogène (Checkoway et coll., 1989). Elles permettent l'étude d'un large spectre de maladies dans la même population, et ont joué un rôle important depuis le milieu du 20e siècle dans la mise en évidence de la cancérogénicité chez l'humain d'agents et de procédés industriels (Axelson, 1979; Breslow et Day, 1987; Checkoway et coll., 2004; Rothman et Greenland, 2008; Attfield et coll., 2012). Elles sont toutefois peu adaptées aux maladies plus rares, incluant plusieurs sites de cancer, qui requièrent une grande taille d'échantillon et un suivi prolongé pour obtenir une puissance statistique acceptable (Siemiatycki et coll., 1981; Blair et coll., 1996; Siemiatycki et coll., 2006; Rothman et Greenland, 2008). Cette limite peut toutefois être partiellement surmontée en combinant plusieurs cohortes (p.ex. Steenland et coll., 2001; Daniels et coll., 2014). Par ailleurs, les études de cohorte sont souvent peu outillées pour tenir compte des facteurs indirects pouvant confondre l'association entre l'exposition et les maladies, hormis certains paramètres tels le sexe, l'âge ou l'origine ethnique des sujets, puisque l'information provient généralement de données administratives (p. ex. registres d'emploi)

(Blair et coll., 1996; Siemiatycki et coll., 2006). Le choix d'une population de référence représentant le groupe de comparaison constitue également un aspect critique de ce type d'étude, en raison notamment de l'effet du travailleur en bonne santé (*healthy worker effect*) (Li et Sung, 1999) pouvant mener à une sous-estimation du risque.

De leur côté, les études cas-témoins pallient à l'utilité limitée des cohortes industrielles dans l'étude des maladies « rares » en associant un groupe de sujets ayant la maladie d'intérêt, les « cas », à des sujets exempts de cette maladie formant le groupe témoin. Les témoins peuvent être recrutés par un échantillonnage dans la population générale, ou parmi des patients atteints d'une autre maladie (Blair et coll., 1996; Checkoway et coll., 2007). Les études cas-témoins permettent l'étude d'un spectre plus large de facteurs de risque en lien avec la maladie, reliés par exemple à l'environnement de travail, à l'environnement général, aux habitudes de vies ou à des traits génétiques. Le nombre généralement plus faible de sujets en comparaison aux cohortes industrielles facilite d'autant plus la collecte d'informations plus détaillées et diversifiées (Pearce et coll., 1989). De plus, leur caractère rétrospectif rend leur recrutement et collecte relativement courts. En contrepartie, l'évaluation de l'exposition professionnelle est particulièrement complexe puisque la carrière des sujets couvre une multitude de milieux de travail, comparativement aux études de cohortes couvrant un secteur précis.

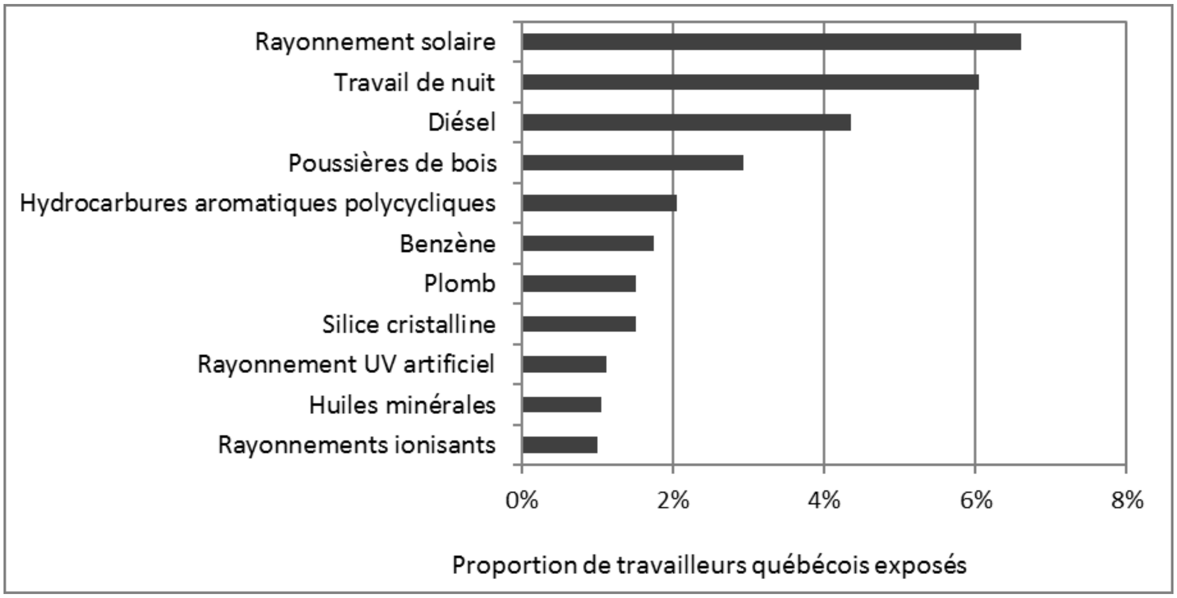
1.1.2 Problématique de l'évaluation de l'exposition dans les études cas-témoins

Les études de cohortes industrielles ne couvrent généralement qu'un nombre restreint de milieux de travail et d'agents, représentant une population d'étude plus homogène relativement à la population générale. Des données administratives et des mesures d'hygiène industrielles peuvent également être disponibles, facilitant l'estimation de l'exposition des travailleurs

(Stewart et coll., 1996; Nieuwenhuijsen et coll., 2006; Sahmel et coll., 2010). À titre d'exemple, plus de 2 000 000 mesures historiques de silice cristalline, remontant jusqu'aux années 1920, étaient disponibles pour une étude regroupant 10 cohortes industrielles (Steenland et coll., 2001; t' Mannetje et coll., 2002). Des conditions historiques peuvent aussi être simulées puis mesurées pour caractériser l'exposition passée (Nieuwenhuijsen et coll., 2006; Sahmel et coll., 2010).

En contrepartie, l'historique professionnel des sujets recrutés dans les études cas-témoins de population englobe un large spectre de professions, d'industries et d'agents sur une période de temps s'échelonnant parfois sur plusieurs décennies. Cette multiplicité de circonstances professionnelles représente un handicap à la collecte d'informations qualitatives ou quantitatives qui sont spécifiques au milieu de travail des sujets (Siemiatycki, 1996). De plus, bien que ces études permettent en théorie l'examen d'une plus grande diversité de facteurs de risque, et donc de contaminants retrouvés en milieu de travail, la proportion de sujets exposés à chacun des agents – ou prévalence de l'exposition – est relativement faible, à l'instar de la population (Blair et coll., 1996; McGuire et coll., 1998; Fritschi et coll., 2009). À titre d'exemple, le rayonnement solaire représentait l'agent le plus prévalent en milieu de travail au Québec selon une évaluation réalisée par Labrèche et coll. (2012), avec une proportion de travailleurs exposés estimée à 6,6% (Figure 1). Une faible prévalence pénalise d'autant plus les erreurs dans la classification de l'exposition des sujets, car elle mène à une plus grande sous-estimation des associations observées entre un agent et la maladie et à une perte de puissance statistique lorsque ces erreurs sont réparties de manière comparable entre les cas et les témoins (Flegal et coll., 1986; Armstrong, 2003).

Figure 1. Proportion de travailleurs québécois exposés aux dix cancérogènes les plus fréquents, tous secteurs d'activité confondus



Adapté de Labrèche et coll. (2012)

Finalement, la rareté des mesures historiques d'hygiène provenant des milieux de travail des sujets a traditionnellement impliqué l'usage d'estimations semi-quantitatives de l'exposition (p. ex. faible, moyenne, élevée). Or, les estimations quantitatives sont généralement préférables aux indices semi-quantitatifs pour dériver des relations dose-réponse entre une agent et la maladie d'intérêt (Berglund et coll., 2001; Sahmel et coll., 2010). En somme, l'évaluation de l'exposition rétrospective dans les études cas-témoins sur le cancer requiert de caractériser le plus fidèlement possible et à partir d'informations limitées, l'exposition d'un grand nombre de sujets au travers d'une multitude de milieux professionnels et sur une période s'étalant sur plusieurs décennies. Les principales approches utilisées pour estimer l'exposition passée dans ces études sont détaillées à la section suivante.

1.2 Méthodes traditionnelles d'évaluation de l'exposition dans les études cas-témoins

1.2.1 Titres d'emploi ou secteurs d'activité économiques

L'estimation du risque d'une maladie associé à un titre d'emploi ou secteur industriel représente une approche généralement rapide et peu coûteuse (Goldberg et Hémon, 1993; Teschke, 2003). La validité des informations fournies par les sujets sur les emplois occupés durant leur carrière lors d'entrevues est généralement reconnue comme très bonne, bien qu'elle puisse être plus faible pour des emplois plus anciens ou de courte durée, ou pour un historique professionnel plus complexe (Baumgarten et coll., 1983; Bourbonnais et coll., 1988; McGuire et coll., 1998; Teschke et coll., 2002). La limite majeure de ce type d'analyse est qu'elle ne permet pas de mettre en évidence les agents spécifiques potentiellement impliqués dans les associations observées. À titre d'exemple, un risque de cancer de la prostate accru chez les sujets ayant œuvré comme peintres a été observé dans une étude montréalaise récente, présentée à l'Annexe 1 (Sauvé et coll., 2016). Or, les activités de peinture peuvent impliquer des expositions à plusieurs types de substances, dont des solvants, des pigments, des vernis et des produits décapants. Puisque l'exposition à une substance peut être distribuée à travers plusieurs professions (Vineis et Blair, 1992; Teschke, 2003; Siemiatycki et coll., 2006), le regroupement des sujets fondé sur la nature des expositions, plutôt que sur la profession, permet non seulement de mettre en évidence les agents responsables, mais aussi d'augmenter la puissance statistique d'une étude pour détecter des associations (Siemiatycki, 1991).

Tableau I : Exemples d'emplois et secteurs industriels évalués par le Centre international de recherche sur le cancer

Titres d'emplois	Groupe CIRC[1]	Secteurs et procédés industriels	Groupe CIRC
Peintres	1	Fonderies de fer et d'acier	1
Barbiers et coiffeurs	2A	Pavage et couvrage de toiture	2A
Charpentiers et menuisiers	2B	Nettoyage à sec	2B
Pompiers	2B	Fabrication de peinture	3

1. Agent cancérogène pour l'homme; 2A. Agent probablement cancérogène; 2B. Agent peut être cancérogène; 3. Inclassable. Source : Centre international de recherche sur le cancer (2017)

L'évaluation fondée sur l'emploi ou l'industrie peut toutefois être utile dans le cas d'exposition à des mélanges complexes ou à des agents inconnus (Teschke, 2003; Vermeulen, 2016), ou dans les cas où l'exposition à un agent est principalement limitée à un emploi particulier tel l'exemple historique de l'exposition à la suie chez les ramoneurs. Le Centre international de recherche sur le cancer (CIRC) a également évalué le risque de cancer associé à des titres d'emploi et de secteurs et procédés industriels, dont quelques exemples sont présentés au Tableau I. Certaines des professions évaluées tels les peintres, l'industrie des fonderies du fer et de l'acier et les mines d'hématites souterraines, sont associées à un potentiel cancérogène reconnu. Néanmoins, l'utilité des titres d'emploi ou d'industrie est principalement limitée à la génération d'hypothèses aiguillant vers de nouvelles pistes de recherche, et représentent un point de départ vers des méthodes permettant d'évaluer l'exposition à des agents spécifiques.

1.2.2 Auto-évaluation de l'exposition par les sujets

Afin d'évaluer l'exposition à des agents spécifiques, les sujets peuvent être sollicités quant aux agents auxquels ils auraient pu être en contact durant leur carrière à l'aide de grilles d'auto-évaluation (Benke et coll., 2001; Teschke, 2003; Neilson et coll., 2007). La validité de l'information rapportée peut toutefois être influencée par la nature des agents et aux connaissances, aux perceptions et à la mémoire des sujets. Par exemple, l'évaluation peut être moins bonne pour les substances qui ne peuvent pas être détectées par les sens (Gérin et Siemiatycki, 1991; Teschke et coll., 2002); de plus, la perception sensorielle peut varier en fonction de facteurs tels l'âge et le tabagisme (pour l'odorat) (Stewart et Stenzel, 1999). La perception ou le rappel des expositions peut également être influencée par la durée et l'ancienneté des emplois, par le sexe et le type de répondant (sujet ou tierce personne) (Benke et coll., 2001; Neilson et coll., 2007; Quinn, 2011). Finalement, le rappel des expositions passées peut différer entre les cas et les témoins et potentiellement mener à une surestimation du risque (Teschke, 2003; de Vocht et coll., 2005), bien que ce phénomène n'ait pas nécessairement été observé dans toutes les études (Neilson et coll., 2007; Hardt et coll., 2014). Compte tenu de ses limites, l'auto-évaluation ne devrait préférablement pas constituer l'unique source d'information pour estimer l'exposition rétrospective (McGuire et coll., 1998).

1.2.3 Matrices emplois-expositions

Une matrice emplois-expositions (MEE) représente un tableau croisé doté d'un axe déclinant par exemple une série de professions, d'industries ou de tâches, et d'un autre axe défini par une liste d'agents. Chaque combinaison unique de ces deux axes, ou cellule, contient une estimation de l'exposition. Certaines MEE comportent un troisième axe pour la période temporelle, permettant de moduler les estimations de l'exposition à un agent dans temps au sein des professions ou industries. Les MEE permettent ainsi d'assigner des expositions à une liste agents pour tous les emplois dans la population d'étude pour lesquels le titre d'emploi ou de secteur d'activité est disponible.

Le développement des MEE en hygiène du travail et dans les études de cohortes industrielles remonte au milieu du 20e siècle (Rappaport, 2009). Ce n'est toutefois que depuis les années 1980 que des MEE dites « générales », couvrant l'ensemble des professions ou secteurs d'activité dans une population, ont fait leur apparition. La MEE développée par Hoar et coll. dans les années 1980 (1980; 1983) aux États-Unis, permettait de relier approximativement 500 combinaisons d'emplois et d'industries à une estimation semi-quantitative (faible, modérée et élevée) de l'exposition à une série de cancérogènes avérés ou soupçonnés, pour un total de près de 15 000 correspondances emploi-exposition. À la même époque, Pannett et coll. (1985) ont développé une MEE générale pour le Royaume-Uni, couvrant l'exposition à 49 agents pour près de 700 catégories d'emploi. La population des cellules d'une MEE par des estimations de l'exposition repose traditionnellement sur le jugement d'experts (p.ex. hygiénistes, médecins du travail, toxicologies), alimenté par diverses sources d'information telles la littérature, des mesures d'exposition en milieu de travail, et des MEE existantes (Kromhout et Vermeulen, 2001).

L'élaboration de la matrice FINJEM (Kauppinen et coll., 1998) par des chercheurs de l'institut finnois de santé au travail dans les années 1990 représente une évolution dans les MEE générales. FINJEM contient trois axes, soit un axe pour les professions selon une classification finlandaise (N=311), un axe d'agents (N=84), incluant des facteurs physiques et psychosociaux, et un axe temporel divisé en 9 périodes remontant à 1945 (Kauppinen et coll., 2014). Chaque cellule contient une estimation de la prévalence, ou probabilité d'exposition (exprimée en pourcentage) et de l'intensité de l'exposition, cette dernière étant exprimée sur une échelle quantitative (p. ex. concentration en partie par millions) pour la majorité des agents. Les estimations ont été développées par expertise, alimentée par une banque de données de mesures d'hygiène (Kauppinen et coll., 1998). Depuis sa création, FINJEM a été appliquée à l'évaluation de l'exposition rétrospective dans les études étiologiques et pour la surveillance des expositions et maladies professionnelles dans la population (Kauppinen et coll., 2014). Elle a également servi de ressource pour le développement de MEE adaptées à d'autres populations (Kauppinen et coll., 2009; t' Mannetje et coll., 2011; García et coll., 2013; van Tongeren et coll., 2013).

Les MEE du programme français MATGÉNÉ (Févotte et coll., 2011), visant à développer des MEE pour la population française, ont un format semblable à la matrice FINJEM, avec une estimation de la prévalence et de l'intensité pour chaque cellule. Il s'agit d'une série de MEE spécifiques à un agent (p.ex. poussières de farine) ou à un groupe d'agents (p. ex. solvants chlorés). Ces matrices sont également destinées à des usages multiples tels l'évaluation de l'exposition professionnelle rétrospective dans les études cas-témoins (Luce et coll., 2011) et la surveillance de l'exposition dans la population (Luce et Févotte, 2006).

La force des MEE est qu'elles permettent d'assigner automatiquement une ou plusieurs estimations de l'exposition à un sujet sur la seule base de la profession ou du secteur industriel. Elles permettent ainsi d'évaluer l'exposition professionnelle dans les études de registres de décès ou de cancer pour lesquels l'information sur la profession ou l'industrie est disponible et où la collecte d'éléments descriptifs plus détaillés sur les emplois n'est pas envisageable (p. ex.

Pukkala et coll., 2005; Luce et Févotte, 2006; Kauppinen et coll., 2009). Dans un même ordre d'idées, elles peuvent constituer la seule approche faisable pour de vastes études comportant des dizaines ou des centaines de milliers de sujets, afin de réduire le volume d'information à récolter et à analyser sur les emplois occupés (p. ex. Sadhra et coll., 2016). L'utilisation du titre d'emploi ou d'industrie comme seul déterminant de l'exposition permet également de minimiser les différences potentielles dans l'information rapportée sur les emplois occupés entre les cas et les témoins (Goldberg et coll., 1993). L'automatisation du processus d'estimation de l'exposition par les MEE représente par contre une limite importante de cette méthode, puisque chaque emploi à l'intérieur d'une même profession se voit attribuer la même exposition. Cette approche implique donc des erreurs de classification si le profil d'exposition des emplois au sein d'une profession n'est pas homogène (Siemiatycki et coll., 1989; Dewar et coll., 1991; McGuire et coll., 1998; Teschke et coll., 2002; Burstyn et coll., 2012). Les MEE avec une probabilité d'exposition exprimée en pourcentage permettent de refléter la variabilité dans l'exposition dans une profession, mais leur application requiert de définir un seuil pour assigner un statut d'exposé aux emplois. Par exemple, pour une cellule avec une probabilité d'exposition de 10%, tous les emplois seraient considérés exposés en utilisant un seuil minimal de 5%, et seraient tous non-exposés avec un seuil plus strict (et plus spécifique) de 95%. En pratique, des seuils intermédiaires entre 25% et 50% sont généralement utilisés comme compromis (Zheng et coll., 2005; Peters et coll., 2011a; Lacourt et coll., 2013) et l'impact du seuil choisi peut faire l'objet d'analyses de sensibilité (Burstyn et coll., 2012; Lacourt et coll., 2013). Cette limite peut être réduite en définissant des groupes plus précis, mais elle implique toutefois une augmentation du nombre de cellules pour lesquelles l'exposition doit être estimée (Bouyer et Hémon, 1993; Plato et Steineck, 1993).

Finalement, l’extrapolation des estimations d’une MEE développée pour une population à une autre peut être limitée par des problèmes de compatibilité entre les systèmes de classification des emplois, ce qui peut également nécessiter un recodage des emplois vers un autre système et représenter une autre source d’erreurs de classification, et par des différences dans la composition industrielle entre les régions qui peuvent influencer la prévalence des expositions (t' Mannetje et coll., 2011; Lavoué et coll., 2012c; Koeman et coll., 2013). L’évolution du marché du travail, et la révision périodique des systèmes de classification pour tenir compte de cette évolution, peut aussi limiter l’utilité des MEE historiques dans l’établissement de portraits actuels de l’exposition dans la population.

1.2.4 Évaluation sur une base individuelle par panels d’experts

Les limites associées aux titres d’emploi et à l’autoévaluation par les sujets a mené au développement d’une approche d’évaluation de l’exposition rétrospective par expertise lors d’une vaste étude montréalaise portant sur de multiples sites de cancer, menée dans les années 1980 (Siemiatycki et coll., 1981; Gérin et coll., 1985; Siemiatycki, 1991). L’approche par expertise est fondée sur le principe que les experts, par leur formation et leur expérience, possèdent une meilleure connaissance des expositions comparativement aux sujets eux-mêmes, et peuvent ainsi fournir des estimations plus fiables (Kromhout et coll., 1987; Teschke et coll., 2002).

L’approche a comme point de départ la collecte, lors d’entrevues, de descriptions d’emplois (p. ex. profession, tâches, procédés), qui couvrent l’ensemble de la carrière des sujets pour les études rétrospectives, à l’aide de questionnaires professionnels généraux. Des questionnaires spécifiques à certaines professions ou industries peuvent également être utilisés pour obtenir de l'information plus détaillée sur des tâches et activités spécifiques plus complexes (p.ex. procédés de soudage). Ces informations sont ensuite revues par un ou plusieurs experts pour estimer l'exposition à une liste d'agents qui est généralement représentée par des indices semi-quantitatifs. L'expérience des experts est mise à profit pour traduire les descriptions d'emploi en estimations de l'exposition. Cette expérience est appuyée par des notes personnelles, la littérature scientifique et des ouvrages couvrant les domaines tels la chimie, la médecine et l'hygiène du travail et les procédés industriels (Goldberg et coll., 1986; Siemiatycki, 1991), et plus récemment l'internet (Fritschi et coll., 2003). Des consultations avec des experts externes et des visites industrielles peuvent également complémenter la collecte d'information (Gérin et coll., 1985; Sahmel et coll., 2010). Des MEE peuvent également représenter une source d'information additionnelle (Semple et coll., 2004; Purdue et coll., 2017).

Contrairement aux MEE, cette approche permet de prendre en compte des facteurs spécifiques à chaque sujet pouvant influencer l'exposition, tels les tâches, procédés, caractéristiques du milieu de travail, et protection individuelle et collective, dans l'estimation de l'exposition (Smith et coll., 2005; Bhatti et coll., 2011; Offermans et coll., 2012). Ce niveau de détail permet de réduire les erreurs de classification et d'augmenter la puissance statistique comparativement à d'autres approches (Siemiatycki et coll., 1989; Dewar et coll., 1991). Pour cette raison, l'évaluation par expertise a historiquement été considérée comme l'approche la plus rigoureuse dans le cadre d'études rétrospectives populationnelles en l'absence de mesures objectives spécifiques aux emplois occupés par les sujets (Bouyer et Hémon, 1993; McGuire et coll., 1998). L'absence de mesures objectives représente toutefois une limite à l'évaluation de la validité des estimations par expertises (McGuire et coll., 1998; Teschke et coll., 2002).

L'approche par expertise nécessite toutefois un investissement considérable en matière de temps, de ressources humaines et de documentation permettant l'évaluation de l'exposition à un large spectre d'agents pour l'ensemble des descriptions d'emploi. Le nombre d'experts possédant l'expérience et les compétences suffisantes pour accomplir ces tâches demeure très restreint, ce qui représente un autre obstacle à son utilisation (Siemiatycki, 2007). L'augmentation du nombre de sujets, et du nombre d'emplois à évaluer, implique également une augmentation des ressources requises à son application (Goldberg et Imbernon, 2002).

Les expositions assignées pour un même emploi peuvent également varier entre deux experts ou pour un même expert, selon l'expérience et l'interprétation des informations fournies par les sujets. Les estimations peuvent également diverger au cours de la conduite de l'étude, par exemple en raison de l'accumulation de nouvelles sources d'information (McGuire et coll., 1998; Friesen et coll., 2013). L'évaluation de la cohérence des estimations inter et intra-experts, ainsi que l'utilisation d'une approche par consensus ou par agrégation des estimations de différents experts, permet de quantifier et de minimiser cette variabilité (Goldberg et coll., 1986; Siemiatycki, 1991; Friesen et coll., 2011). Le processus décisionnel entourant l'assignation des estimations représente aussi une sorte de « boîte noire » (Gomez et coll., 1994; Kauppinen, 1996; Stewart et coll., 1996; Stewart et Stenzel, 1999; Kromhout, 2002) dont l'exploration a fait l'objet de recherches récentes pour accroître la transparence et la reproductibilité l'approche par expertise. En amont, la validité de l'expertise repose également sur la qualité et l'exhaustivité des informations rapportées par les sujets, qui peuvent être plus faibles si ces derniers ne sont pas en mesure de décrire précisément les activités ou produits utilisés ou si celles-ci ont été obtenues par des questions vagues ou par des intervieweurs moins familiers avec les principes de l'hygiène du travail (Stewart et Stewart, 1994). Finalement, puisque l'évaluation de l'exposition porte uniquement sur l'échantillon de sujets spécifiques à une étude, les informations sur l'exposition ne peuvent pas directement être appliquées à d'autres sujets, comparativement aux MEE dont le potentiel de réemploi constitue un avantage marqué (Hoar, 1983). Des efforts d'exploitation de données d'évaluation existantes afin de permettre leur réutilisation sont en cours et présentés à la section 1.3.2.

1.3 Méthodes récentes d'évaluation de l'exposition rétrospective

Au cours des 10 dernières années, certaines approches ont été mises de l'avant pour améliorer l'efficacité de l'évaluation de l'exposition en combinant l'automatisation du processus permis par les MEE à la précision associée à la méthode par expertise. Ces approches sont basées d'une part sur le développement d'algorithmes de classification, et d'autre part, sur l'exploitation des expertises réalisées au cours d'études antérieures. Des efforts ont également été consacrés pour intégrer des estimations quantitatives, associées principalement à des mesures dans l'air mais également dans les matrices biologiques, dans les études de population à dimension rétrospective (p. ex. Bosch de Basea et coll., 2011; DellaValle et coll., 2015).

1.3.1 Algorithmes décisionnels

Bien qu'elle représente l'approche de référence, la revue de chaque description d'emploi par des experts pour évaluer les expositions représente un processus lourd et complexe. La constance dans les estimations peut également être problématique, bien qu'elle puisse être contrôlée par des principes directeurs pour guider les évaluations, par exemple par un classement des professions par catégorie d'exposition (Siemiatycki et coll., 1991) ou par l'évaluation d'un échantillon d'emplois pour calibrer les évaluations (Rocheleau et coll., 2011).

Un développement récent consiste en l'élaboration de règles de décisions prédéterminées par des experts pour associer les circonstances professionnelles à des estimations de l'exposition (Fritschi et coll., 2009; Pronk et coll., 2012). Cette approche permet de faciliter l'évaluation de l'exposition en simplifiant la collecte d'information en ciblant les questionnaires et entrevues aux situations avec un potentiel d'exposition (Fritschi et coll., 2009). L'application des algorithmes aux réponses des questionnaires permet aussi d'assigner automatiquement des estimations de l'exposition, permettant ainsi un gain de temps appréciable comparativement à l'approche traditionnelle par expertise (Fritschi et coll., 2009; Fritschi et coll., 2012), tout en augmentant la cohérence et la reproductibilité des évaluations (Pronk et coll., 2012). Les règles de décision permettent la prise en compte de facteurs individuels dans l'attribution des niveaux d'exposition, en comparaison aux MEE basées uniquement sur la profession comme seul déterminant. L'approche par règles de décision a notamment été appliquée pour estimer l'exposition passée pour des travailleurs atteints de maladies reliées à l'amiante à des fins d'indemnisation (Macfarlane et coll., 2012), et pour dresser des portraits de l'exposition professionnelle dans la population australienne (p. ex. Carey et coll., 2014; Driscoll et coll., 2016; Si et coll., 2016).

Comparativement à l'approche par expertise traditionnelle, la rigidité des règles de décision implique une standardisation de la collecte d'information laissant moins de latitude pour explorer des circonstances professionnelles moins fréquentes associées à l'exposition. L'application de méthodes d'exploration de texte (« text mining ») représente une stratégie en développement pour rendre compatible l'emploi de questions ouvertes aux algorithmes (Friesen et coll., 2014). La performance des algorithmes est également moins grande pour des situations d'exposition intermédiaires ou incertaines comparativement à une évaluation par expertise individuelle (Pronk et coll., 2012; Friesen et coll., 2013). Les algorithmes peuvent toutefois permettre de trier les emplois pour identifier ces situations et permettre aux experts de concentrer leurs efforts sur des cas plus incertains (Fritschi et coll., 2012). Finalement, une expertise spécialisée couplée à un bassin de documentation est requise afin de permettre l'élaboration des algorithmes. L'utilisation de MEE développées pour des cohortes industrielles a été envisagée comme source d'information pour le faciliter le développement des règles de décision (Behrens et coll., 2012).

1.3.2 Exploitation d'expertises passées

Une autre stratégie mise de l'avant dans l'objectif d'améliorer l'efficacité et la précision de l'évaluation de l'exposition consiste à exploiter l'information contenue dans les expertises réalisées lors d'études antérieures. Les expertises d'emplois individuels par profession peuvent par exemple être agrégées par profession afin de présenter l'information sur l'exposition sous forme de MEE. Par exemple, la synthèse par professions d'évaluations provenant d'une étude cas-témoins sur le cancer du poumon en Europe de l'Est a constitué une source d'information dans l'élaboration d'une MEE générale pour la Nouvelle-Zélande (t' Mannetje et coll., 2011).

Plus près de nous, les expertises réalisées dans les études cas-témoins montréalaises sont à la source d'une approche par expertise hybride d'évaluation de l'exposition. Dans cette approche, des profils d'exposition par professions construits à partir des expertises existantes, ont servi de source d'information pour les experts afin de faciliter l'évaluation de l'exposition à plus de 300 agents dans une étude montréalaise sur le cancer de la prostate (étude PROtEuS) comptant approximativement 4000 sujets (Blanc-Lapierre et coll., 2015; Sauvé et coll., 2016). Par ailleurs, les données montréalaises ont également été utilisées pour orienter l'adaptation de la matrice

FINJEM à une population plus large couvrant sept pays, dans le cadre de l'étude multicentrique INTEROCC sur le cancer du cerveau (Lavoué et coll., 2012c; van Tongeren et coll., 2013).

Les expertises passées ont par ailleurs été exploitées en modélisant les expositions assignées par les experts par les informations rapportées par les sujets dans les questionnaires, à l'aide de réseaux de neurones artificiels (Black et coll., 2004) ou de modèles d'arbres de classification et de régression (Wheeler et coll., 2013; Wheeler et coll., 2015). Ces exercices visaient à identifier quels étaient les déterminants principaux dans les descriptions d'emplois (p.ex. profession, tâches ou équipements) associés à la présence de l'exposition à un agent et à son niveau d'intensité et de fréquence, dans une optique de caractériser le processus décisionnel interne des experts et d'augmenter la reproductibilité de l'approche. L'application de ces modèles permet de mettre en lumière les sources de désaccords entre les experts dans les évaluations pour améliorer la cohérence dans les évaluations, et de raffiner les questionnaires en tenant compte des éléments les plus fortement associés à l'exposition. Finalement, les paramètres des modèles représentent des règles de décision qui peuvent être appliquées pour estimer l'exposition.

1.3.3 Approches quantitatives dans les études de population

Récemment, des méthodes ont été développées pour permettre d'associer des niveaux quantitatifs de l'exposition aux emplois occupés par les sujets dans les études cas-témoins en combinant une MEE à des mesures d'hygiène industrielle (p. ex. Friesen et coll., 2012; Peters et coll., 2016). Ces avancées s'inscrivent dans un contexte d'exploitation accrue des données d'exposition disponibles dans la littérature et dans les banques de données administratives en hygiène et en épidémiologie professionnelle.

1.3.3.1 Sources de mesures d'exposition

Les articles de périodiques scientifiques représentent une source potentiellement importante de mesures objectives, sans compter les rapports de surveillance environnementale d'organismes de recherche en santé au travail, notamment les *Health Hazard Evaluations* du NIOSH[1] américain (Froines et coll., 1989; Hein et coll., 2010) et de la littérature dite « grise », non publiée (Burstyn et coll., 2000). Une synthèse de cette littérature permet entre autres d'identifier les sources et déterminants de l'exposition dans le développement de questionnaires, et pour guider l'estimation des niveaux d'exposition dans les MEE ou par expertise (Teschke et coll., 2002).

Les banque de données d'exposition professionnelle administratives, contenant des résultats d'échantillonnages en milieu de travail, forment une autre source d'information objective sur les niveaux d'exposition. Plusieurs pays possèdent des systèmes informatisés d'enregistrement de mesures de surveillance environnementale couvrant un large spectre de secteurs d'activités économiques et de contaminants (Stewart, 1999; Teschke et coll., 2002; Lavoué, 2006; ter Burg, 2014). Par exemple, la banque IMIS (*Integrated management information system*) contient plus de 1,5 million de résultats d'échantillonnages réalisés par des inspecteurs de l'organisme réglementaire fédéral américain OSHA[2] depuis 1972. IMIS a été identifiée il y a plus d'un quart de siècle comme une source potentielle d'information sur l'exposition rétrospective pour les études en épidémiologie professionnelle (Froines et coll., 1989; Stewart et Rice, 1990; Lavoué et coll., 2012a). D'autres banques majeures incluent COLCHIC en France (Vincent et Jeandel, 2001), qui a notamment servi de source d'information dans la surveillance de l'exposition à des

[1] *National Institute of Occupational Safety and Health*
[2] *Occupation Safety and Health Administration*

cancérogènes du programme européen CAREX (Vincent et coll., 1999); la banque allemande MEGA (Gabriel et coll., 2010); la banque anglaise NEDB (Burns et Beaumont, 1989) et la banque italienne SIREP (Scarselli et coll., 2007). Plus près de nous, la banque LIMS, de l'Institut recherche Robert Sauvé en santé et en sécurité du travail, contient près de 600 000 mesures sur la période 1985-2008 (Lavoué et coll., 2012b). La banque canadienne CWED (*Canadian Workplace Exposure Database*) (Hall et coll., 2011), comptant approximativement 500 000 mesures, résulte d'un projet de mise en commun de banques de données existantes d'organismes provinciaux et territoriaux en santé au travail, et dont une description est présentée à la section 2.1.3.

Les questionnements quant à la représentativité des mesures contenues dans les banques de données d'exposition, associées principalement à l'échantillonnage non-aléatoire des postes et milieux de travail ciblant entre autres les situations à risque de surexposition (Stewart et Rice, 1990; Lavoué, 2006; Viet et coll., 2008), ont formé une limite à leur utilisation dans les études de population. Ces banques ne couvrent pas nécessairement l'ensemble des industries ou professions et de la période temporelle d'une population d'étude. De plus, puisque les banques de données d'exposition contiennent des mesures visant généralement à documenter la conformité réglementaire des entreprises, certains agents non couverts par la législation peuvent donc en être absents (Froines et coll., 1989). Finalement, les informations contextuelles et les titres d'emplois permettant d'interpréter les niveaux d'exposition peuvent être manquantes ou de qualité inégale, ce qui peut représenter un obstacle à leur interprétation et à leur utilité dans le cadre d'études épidémiologiques et d'analyses du risque (Tielemans et coll., 2002).

1.3.3.2 Intégration d'estimations quantitatives dans les études de population

Les approches récentes permettant l'estimation quantitative de l'exposition pour des études de population, initiées par Wild et coll. (2002), sont basées sur la combinaison d'une MEE semi-quantitative et des mesures d'une banque de données afin d'estimer des concentrations moyennes associés à chaque catégorie semi-quantitative, permettant au final le développement d'une MEE entièrement quantitative. Cette approche permet de tirer profit des forces des deux sources de données. D'une part, la MEE permet de couvrir l'ensemble des professions dans la population, et les niveaux semi-quantitatifs des cellules fournissent une information sur la présence de l'exposition ainsi qu'un jugement semi-quantitatif sur son amplitude. D'autre part, les mesures contenues dans la banque de données ne couvrent pas nécessairement l'ensemble des professions, mais elles permettent une estimation quantitative pour les situations où l'exposition est présente et a été quantifiée. L'estimation de niveaux quantitatifs aux cellules de la MEE consiste à ajuster un modèle linéaire mixte avec les concentrations comme variable dépendante, et pour les variables prédictives, la catégorie d'exposition pour la profession dans la MEE est entrée comme effet fixe, alors que la profession elle-même comme effet aléatoire. L'estimation d'une concentration moyenne pour chaque catégorie d'intensité permet ainsi d'attribuer un niveau d'exposition quantitatif pour n'importe quelle profession dans la MEE. Pour les professions représentées dans la banque de données, le niveau d'exposition dans la MEE quantitative combine la concentration moyenne associée à la catégorie semi-quantitative, et la concentration moyenne dans les mesures spécifiques à cette profession. Pour les professions non-représentées dans la BEDP, l'estimation dans la MEE quantitative est uniquement basée sur la concentration moyenne associée à la catégorie semi-quantitative.

Ce cadre conceptuel a été appliqué à deux grandes études de population, l'une portant sur l'exposition à la silice cristalline dans une étude multicentrique combinant 14 études cas-témoins sur le cancer provenant de 13 pays européens et du Canada (Peters et coll., 2011b). Une banque contenant près de 24 000 mesures, remontant jusqu'au milieu des années 1970, a permis d'assigner des niveaux quantitatifs à une MEE générale, tout en tenant compte de facteurs tels l'année, la durée de mesure et la région dans les estimations. L'approche de modélisation développée pour la silice cristalline a récemment été étendue à l'amiante, au chrome hexavalent, au nickel et au benzo(a)pyrène (Peters et coll., 2016). L'autre application a porté sur une étude de cohorte chinoise comptant près de 75 000 femmes, pour laquelle une banque de données comprenant approximativement 71 000 mesures d'exposition au benzène était disponible (Friesen et coll., 2012). Les professions dans la MEE étaient stratifiées par industrie, permettant d'estimer des niveaux d'exposition différents entre les emplois au sein d'une même profession (p.ex. peintres) en tenant compte de l'industrie (p.ex. peintres dans l'industrie maritime vs peintres dans l'industrie du verre). La méthode a ensuite été reprise dans le développement d'estimations de l'exposition aux poussières et fumées contenant du plomb (Koh et coll., 2014).

1.3.4 Biomarqueurs d'exposition

Les mesures quantitatives de l'exposition ne sont pas exclusivement limitées à la dose dite externe. À ce titre, les biomarqueurs permettent d'identifier et/ou de caractériser l'exposition *a posteriori* à des xénobiotiques par des mesures effectuées dans les matrices biologiques (p. ex. sang, urine) (Paustenbach, 2000). Les biomarqueurs comportent plusieurs avantages. Ils permettent de prendre en compte l'exposition provenant de plusieurs voies (Rappaport et coll., 1995), facilitant l'évaluation lorsque la source est difficile à définir (Bencko, 2011), et d'intégrer l'exposition professionnelle et environnementale. Finalement, les biomarqueurs peuvent être plus représentatifs de l'exposition reçue avec une protection individuelle (p. ex. masques filtrants) comparativement à des mesures dans l'air (Nieuwenhuijsen et Droz, 2003).

L'application des biomarqueurs à l'évaluation de l'exposition rétrospective pour des études sur les causes professionnelles du cancer demeure toutefois restreinte aux agents ayant un longue demi-vie d'élimination tels certains métaux (Lin et coll., 2005), l'arsenic (Bencko, 2011) et des composés organiques persistants (Verner et coll., 2011; DellaValle et coll., 2015). Le nombre de biomarqueurs validés est également limité, et la reconstruction de la dose d'exposition requiert une connaissance élaborée du métabolisme et des mécanismes impliqués (Rappaport et coll., 1995; García et Checkoway, 2003; Checkoway et coll., 2004) qui peuvent aussi être affectés par la maladie et/ou son traitement (Siemiatycki et coll., 2006). Finalement, les coûts associés au prélèvement et à l'analyse des échantillons peuvent être onéreux.

1.4 Résumé de la revue de littérature

L'évaluation par expertise et les MEE ont permis, depuis l'aube des années 1980, de dépasser le seul titre de profession ou de secteur industriel comme unité d'analyse et de permettre l'évaluation d'associations entre le risque de maladie et de multiples expositions chimiques et physiques dans les études de population. Ces méthodes ont ainsi permis une meilleure caractérisation des risques à la santé associés à des agents spécifiques rencontrés dans l'environnement de travail et dans l'environnement général.

L'approche par expertise permet un grand niveau de précision dans l'évaluation de l'exposition à un large spectre d'agents, mais sa complexité représente un obstacle important à son application à grande échelle. L'augmentation de la taille des études, ainsi que la part accrue d'emplois contractuels, à court terme ou à temps partiel (Pold, 2001; Galarneau, 2010) entraîne également un volume croissant d'emplois à évaluer qui, combiné à des ressources retreintes, limite la faisabilité de cette méthode. Ces facteurs ont mené à une popularité accrue des MEE, qui peuvent également être appliquées à plusieurs études en dépit des erreurs de classification sur l'exposition introduites. Le nombre de MEE contemporaines permettant l'évaluation de l'exposition à un large éventail de contaminants et adaptées à une diversité de populations et d'applications est toutefois relativement limité, notamment en Amérique du Nord. En parallèle avec l'augmentation de la taille des études, l'identification de risques de plus en plus faibles (Siemiatycki et coll., 2004) nécessite un raffinement constant des méthodes d'évaluation en termes de précision, de sensibilité et de spécificité, qui inclut notamment l'utilisation d'estimations quantitatives de l'exposition.

1.5 Données d'expertises montréalaises

Le groupe de recherche en épidémiologie environnementale et santé des populations à Montréal constitue le berceau de l'approche par expertise individuelle. Depuis son développement dans les années 1980, l'approche a été appliquée dans le cadre de quatre grandes études cas-témoins sur le cancer sur une période de 25 ans, combinant près de 9000 sujets. Pour chaque étude, l'équipe d'experts a évalué l'exposition à environ 300 agents chimiques et physique pour chaque description d'emploi. Tel que présenté au Chapitre 3, la somme des études constitue un bassin d'information sur l'exposition associée à plus de 31 000 emplois sur une période s'échelonnant sur quelques 80 années, et représente 50 personnes-années d'expertise. Cette source de données représente ainsi une ressource documentaire unique au monde, entre autres par la longue liste d'agents évalués qui inclut plusieurs agents peu prévalents et rarement représentés dans les sources d'information actuelles. La banque de données d'expertises montréalaises représente une ressource rarement exploitée jusqu'à maintenant. La valorisation et le partage des expertises montréalaises, ainsi que leur couplage à des mesures objectives d'hygiène industrielle en tirant profit des méthodes quantitatives récentes, s'inscrit dans le contexte plus large de l'amélioration de l'évaluation de l'exposition rétrospective professionnelle dans les études de population.

1.6 Objectifs de la recherche

1.6.1 Objectif général

L'objectif principal de cette recherche vise à développer et à caractériser des outils permettant d'améliorer l'évaluation de l'exposition professionnelle rétrospective dans les études de population, par l'exploitation de la banque de données de résultats d'évaluation par expertise réalisées dans les études cas-témoins montréalaises.

1.6.2 Objectifs spécifiques de la recherche

L'objectif général se décline en quatre volets, soit :

1. L'élaboration d'une matrice emplois-exposition (matrice CANJEM) réalisée par une synthèse descriptive des évaluations d'experts réalisées dans les études montréalaises par profession ou secteur industriel, agent et période temporelle

2. Le développement d'une approche de modélisation permettant de raffiner les estimations de la matrice CANJEM en tenant compte des relations entre les professions

3. Le développement d'une dimension quantitative de la matrice CANJEM par l'intégration de mesures d'hygiène industrielles historiques canadiennes

4. La comparaison d'une approche hybride combinant expertise individuelle et profils d'emplois basés sur les évaluations d'experts montréalaises avec l'approche par expertise traditionnelle.

1.7 Organisation de la thèse

Cette thèse est organisée en sept chapitres. Le présent chapitre offre une description de la problématique de l'évaluation de l'exposition professionnelle dans les études de population, les méthodes d'évaluation utilisées depuis les 35 dernières années et leurs limites, et introduit les objectifs de la recherche. Le chapitre 2 portant sur les aspects méthodologiques du travail met en scène les sources de données utilisées dans les analyses et une brève description des méthodes statistiques utilisées. Les chapitres 3 à 6 formant le corps de la thèse déclinent les quatre manuscrits associés à la réalisation des objectifs spécifiques définis à la section 1.6.2. Une discussion générale des résultats observés dans les chapitres précédents, assortie des principales conclusions tirées des analyses, vient clore cette thèse au chapitre 7. Finalement, l'Annexe 1 présente un manuscrit tiré d'une évaluation du risque de cancer de la prostate par titre professionnel et secteur industriel, réalisé à partir de données de l'étude PROtEuS en marge des travaux présentés au chapitre 6.

Article 1. Development of the CANJEM job exposure matrix: Bayesian modelling of occupational exposures assigned by experts to over 30,000 jobs spanning 1930-2005. Article en préparation, soumis aux co-auteurs en prévision d'une soumission dans un périodique scientifique

Article 2. Development of the CANJEM job exposure matrix: Bayesian modelling of occupational exposures assigned by experts to over 30,000 jobs spanning 1930-2005. Article en préparation, soumis aux co-auteurs en prévision d'une soumission dans un périodique scientifique

Article 3. Development of quantitative estimates of wood dust exposure in a Canadian general population job-exposure matrix based on past expert assessments. Article en préparation, soumis aux co-auteurs en prévision d'une soumission dans un périodique scientifique

Article 4. A hybrid expert approach for retrospective assessment of occupational exposures in a population-based study. Article en préparation, soumis aux co-auteurs en prévision d'une soumission dans un périodique scientifique

Annexe 1. Occupation, industry, and the risk of prostate cancer: a case-control study in Montréal, Canada. Article publié dans *Environmental Health* vol. 15, n° 1, p. 100 (2016)

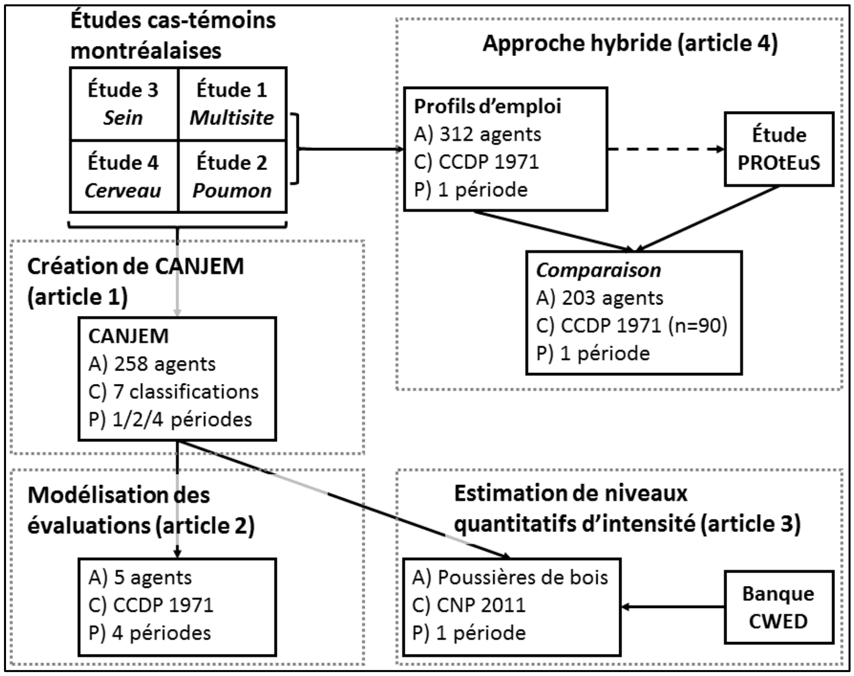
Chapitre 2. Méthodologie

Méthodologie

Les travaux présentés aux chapitres 4 à 7 ont tous impliqué l'utilisation de la banque de données d'exposition compilant les évaluations réalisées dans les quatre grandes études cas-témoins historiques montréalaises. Ce chapitre débute par une brève description de ces études et de la méthode d'évaluation par expertise utilisée. L'étude montréalaise récente PROtEuS sur le cancer de la prostate et l'approche hybride qui y est utilisée pour évaluer l'exposition professionnelle rétrospective, ainsi que la banque de mesures historiques canadienne CWED, sont également présentés. L'organisation de ces sources de données dans les chapitres de la thèse est illustrée à la Figure 1.

La section 2.2 présente un survol de la méthodologie utilisée pour chacun des manuscrits formant les 4 chapitres suivants. Finalement, la section décrit les principes généraux associés aux méthodes statistiques employées, touchant la modélisation des données d'exposition et l'inférence bayésienne.

Figure 1. Cartographie des sources de données selon les articles de la thèse



A : Liste d'agents retenus; C : Classifications professionnelles ou industrielles utilisées (cf. Annexe 2); P) Niveau de résolution pour la période temporelle (cf. section 2.2.1). CCDP : Classification Canadienne Descriptive des Professions; CNP : Classification Nationale des Professions; CWED : Canadian Workplace Exposure Database

2.1 Sources de données d'exposition

2.1.1 Études cas-témoins historiques montréalaises

Les quatre grandes études historiques montréalaises ont été menées entre 1979 et 2004, et ont inclus au total près de 9000 sujets. Le Tableau I présente le nombre de sujets, le nombre d'emplois et la période couverte par l'historique des sujets pour chaque étude, limités aux données d'emploi incluses dans CANJEM, en plus de l'étude PROtEuS décrite à la section 2.1.2. Il est à noter que les données associées à certains sujets inclus dans les études épidémiologiques n'ont pas été utilisées pour la construction de CANJEM (et vice-versa), par exemple pour les sujets n'ayant pas tenu un emploi éligible au cours de leur vie pour cause de maladie. En conséquence, les chiffres rapportés au Tableau I peuvent différer du nombre de sujets rapportés à la section 2.1.1.1.

Tableau I. Résumé des quatre grandes études cas-témoins montréalaises incluses dans CANJEM, et de l'étude cas-témoins récente PROtEuS sur le cancer de la prostate.

Étude	Site	Années de recrutement	Nombre de sujets[1]	Nombre d'emplois[2]	Période couverte par les emplois[2]
1	Multiple	1979-1986	4371	15 067	1920-1985
2	Poumon	1996-2001	2662	10 371	1934-2001
3	Sein	1996-1997	1068	3510	1933-1996
4	Cerveau	2000-2004	659	2725	1956-2004
PROtEuS	Prostate	2005-2009	4013	16 065	1943-2012

1. Pour les études 1-4, les nombres représentent le nombre de sujets dont l'historique d'emploi est inclus dans CANJEM. Pour PROtEuS, il s'agit du nombre de sujets pour lesquels l'historique d'emploi a été obtenu.
2. Pour les études 1-4, les nombres sont basés sur les emplois inclus dans CANJEM. Pour PROtEuS, les nombres correspondent aux emplois éligibles (d'une durée de 2 ans ou plus) évalués par les experts.

2.1.1.1 Description des populations d'étude

Étude multisite

L'étude multisite constitue la première étude réalisée par le groupe, durant laquelle l'approche d'évaluation de l'exposition par expertise a été développée. L'étude a été menée entre 1979 et 1986, et portait sur un total de 19 sites de cancer tels le poumon, la prostate ou la vessie (Siemiatycki et coll., 1987; Siemiatycki, 1991). La population d'étude comprenait un total de 4259 hommes âgés de 35 à 70 ans, répartis entre 3726 cas incidents de cancer (tous sites confondus) recrutés parmi les 18 principaux hôpitaux de la région montréalaise, et 533 témoins recrutés dans la population générale, résidents montréalais et appariés aux cas sur la base de tranches d'âge.

Étude poumon

La deuxième étude, menée de 1996 à 2001 sur des résidents montréalais, comprenait comme série de cas 1203 hommes et femmes âgés de 35 à 75 ans et atteints de tumeurs pulmonaires ou de mésothéliomes. Le groupe de témoins était composé de 1513 sujets recrutés dans la population générale, appariés aux cas selon l'âge, le sexe et la circonscription électorale provinciale (Nkosi et coll., 2012).

Étude sein

L'étude cas-témoins sur le cancer du sein fut menée en 1996 et en 1997 (Labrèche et coll., 2010). 603 femmes âgées entre 50 et 75 ans et diagnostiquées avec une tumeur maligne primaire au sein parmi 18 hôpitaux de la région de Montréal formaient la série de cas. Les témoins étaient quant à eux représentés par 667 femmes diagnostiquées d'un cancer autre que le cancer du sein parmi les mêmes hôpitaux durant la période de recrutement, appariés aux cas selon l'âge.

Étude cerveau

La plus récente étude réalisée par le groupe représente le volet canadien de l'étude internationale INTEROCC (Lacourt et coll., 2013; McLean et coll., 2014) sur les facteurs professionnels associés au cancer du cerveau. Réalisée entre 2000 et 2004 à Montréal, la série de cas regroupait 218 hommes et femmes âgés entre 30 et 59 ans et diagnostiqués avec une tumeur cérébrale (principalement des gliomes ou méningiomes). 414 sujets recrutés dans la population générale, appariés aux cas sur la base de l'âge, du sexe et de la circonscription électorale provinciale, formaient le groupe témoin.

2.1.1.2 Méthodes de collecte de données et d'évaluation de l'exposition

À la suite du recrutement des sujets, des entrevues ont été réalisées afin de récolter des informations sur une large gamme de facteurs potentiellement associés à la maladie d'intérêt. Ceux-ci incluaient des données socio-économiques (p.ex. revenu, niveau d'éducation, origines ancestrales), les antécédents médicaux, les habitudes de vie (p.ex. tabagisme, utilisation de téléphone cellulaire), des variables anthropométriques (p.ex. indice de masse corporelle) et l'historique résidentiel.

Une large part des entrevues et des questionnaires a été consacrée à la collecte de l'historique professionnel complet par des interviewers spécialement formés par une équipe d'hygiénistes industriels. De plus, ces derniers révisaient les histoires professionnelles recueillies afin de clarifier les informations auprès des interviewers, au besoin. Pour certains emplois d'intérêt, ou associés à un profil d'exposition plus complexe, des questionnaires spécialisés ont été utilisés afin de récolter de l'information supplémentaire. À titre d'exemple, le questionnaire pour les emplois de soudeurs couvrait le type de procédé de soudage, les métaux utilisés, le type d'électrode et l'utilisation de divers solvants, dégraisseurs et matériaux ou outils abrasifs, en plus de la fréquence des différentes activités.

L'approche d'évaluation de l'exposition professionnelle rétrospective par expertise, développée dans le cadre de la première étude, est décrite en détail dans Gérin et coll. (1985) et Siemiatycki et coll. (1991). En résumé, des codes professionnels et de secteurs économiques selon des classifications standardisées ont été assignés pour chaque emploi. Les experts ont ensuite évalué la présence d'une ou de plusieurs expositions parmi une liste comprenant approximativement 300 substances chimiques (incluant des mélanges ou des catégories de produits) et agents physiques. La présence d'une exposition était assignée lorsque son niveau était jugé supérieur à celui retrouvé normalement dans l'environnement général. Trois paramètres étaient utilisés pour caractériser l'exposition à chaque agent pour un emploi donné, soit la fiabilité, l'intensité et la fréquence. La fiabilité représentait le niveau de confiance (possible, probable, certain) de l'expert quant à la présence de l'exposition pour l'emploi évalué. L'intensité de l'exposition était quant à elle évaluée selon une échelle semi-quantitative (faible, moyenne, élevée). La fréquence était quant à elle évaluée sur la base d'une semaine typique de travail.

Afin de faciliter et de standardiser l'attribution des niveaux d'intensité, une série d'emplois ou d'activités jugés *a priori* représentatifs de chaque catégorie a été dressée pour certains agents, servant d'échelle de calibration. Par exemple, l'exposition à des poussières d'acier inoxydable était associée à un niveau faible chez les plombiers ou lors de tâches de gravure, à un niveau moyen pour les travailleurs des fonderies, et à un niveau élevé pour les machinistes ou lors d'opérations de meulage ou de polissage de pièces en acier inoxydable. Ces niveaux étaient utilisés à titre indicatif seulement, et pouvaient être modulés en fonction des descriptions d'emploi rapportées par les sujets.

Lors de chaque étude, une première ronde d'encodage était réalisée, où les emplois étaient répartis parmi l'équipe d'experts sur la base des titres d'emploi. Une deuxième ronde d'évaluation était ensuite effectuée, où les expositions attribuées durant la première ronde étaient revues par un expert différent. Afin de minimiser les biais, les experts étaient « aveugles » quant au statut de cas/témoins des sujets tout au cours du processus de codage des emplois et des expositions chimiques. Dans le cas de désaccords dans les évaluations entre deux experts, les encodages finaux étaient attribués sur la base d'un consensus.

2.1.2 Étude PROtEuS et approche hybride d'évaluation de l'exposition

L'étude cas-témoins PROtEuS (*Prostate Cancer and Environment Study*) vise à explorer des associations potentielles entre de multiples facteurs environnementaux, incluant les expositions à des agents en milieu de travail, et le développement et la progression du cancer de la prostate. Elle inclut 1937 hommes diagnostiqués avec un cancer de la prostate entre 2005 et 2009 parmi les principaux hôpitaux francophones de la région de Montréal. Les témoins ont été recrutés en parallèle à partir d'un échantillon aléatoire de la liste électorale francophone du Québec parmi les circonscriptions électorales représentées par les cas, appariés aux cas selon l'âge par strate de 5 ans.

L'approche hybride d'évaluation de l'exposition utilisée dans l'étude PROtEuS constitue une évolution de l'approche par expertise traditionnelle. Les experts avaient accès à une synthèse par profession d'évaluations réalisées dans l'étude poumon, et dans l'étude multisite pour certains agents complémentaires. Pour chaque profession représentée dans les études antérieures, un tableau descriptif présentait le nombre d'emplois exposés à chaque substance et leur distribution par indice de fiabilité, intensité et fréquence, qui pouvaient également être accompagnés de courts commentaires. Une liste de 295 professions plus complexes a aussi fait l'objet d'une revue approfondie par un expert senior lors de la préparation de l'étude PROtEuS, afin d'attribuer des commentaires plus détaillés et/ou spécifiques et d'ainsi bonifier l'information à la disposition des experts. Une échelle de couleurs, représentant un indicateur visuel de la variabilité dans les niveaux d'exposition, était également utilisée pour permettre aux experts de départager aisément les expositions requérant plus ou moins d'attention.

2.1.3 Banque CWED

La banque CWED résulte d'une initiative entamée en 2008 visant à construire une banque de données d'exposition professionnelle historique couvrant la population canadienne. L'élaboration de CWED est associée au projet CAREX Canada portant sur la surveillance de l'exposition aux substances cancérogènes en milieu de travail au pays. CWED a été constituée en regroupant des banques de données existantes d'organismes provinciaux (Colombie-Britannique, Manitoba, Ontario, Québec, Saskatchewan), territoriaux (Yukon) et fédéraux. CWED représente également un effort d'archivage de données devant l'avenir incertain de certaines banques provinciales (Hall et coll., 2011).

CWED recense approximativement 500 000 mesures réparties entre 350 agents et remontant jusqu'aux années 1960. Le monoxyde et le dioxyde de carbone, les poussières non classées autrement, le toluène, le xylène, les composés du plomb et le formaldéhyde font partie des agents les plus fréquemment rencontrés avec plus de 10 000 mesures chacun. CWED a notamment servi de source de données dans l'évaluation de la prévalence de l'exposition à des agents cancérogènes au Canada (Ge et coll., 2013; Peters et coll., 2015), dans l'estimation du fardeau de cancer d'origine professionnelle au pays (Demers et coll., 2014), et dans la réalisation d'un portrait historique de l'exposition aux isocyanates en Ontario et en Colombie-Britannique (Hon et coll., 2017).

Les paramètres descriptifs accompagnant les mesures dans CWED incluent l'année, la province et la source originale des données, la profession selon la Classification nationale des professions

(CNP), édition 2006 (Statistics Canada, 2007), le nom de l'entreprise et le secteur industriel selon le Système de classification industriel de l'Amérique du Nord (SCIAN), édition 2002 (Statistics Canada, 2003). Les paramètres méthodologiques associés à la collecte des échantillons et leur analyse incluent la durée et la zone (respiratoire ou ambiante) de mesure, la stratégie d'échantillonnage, la méthode analytique et sa limite de quantification.

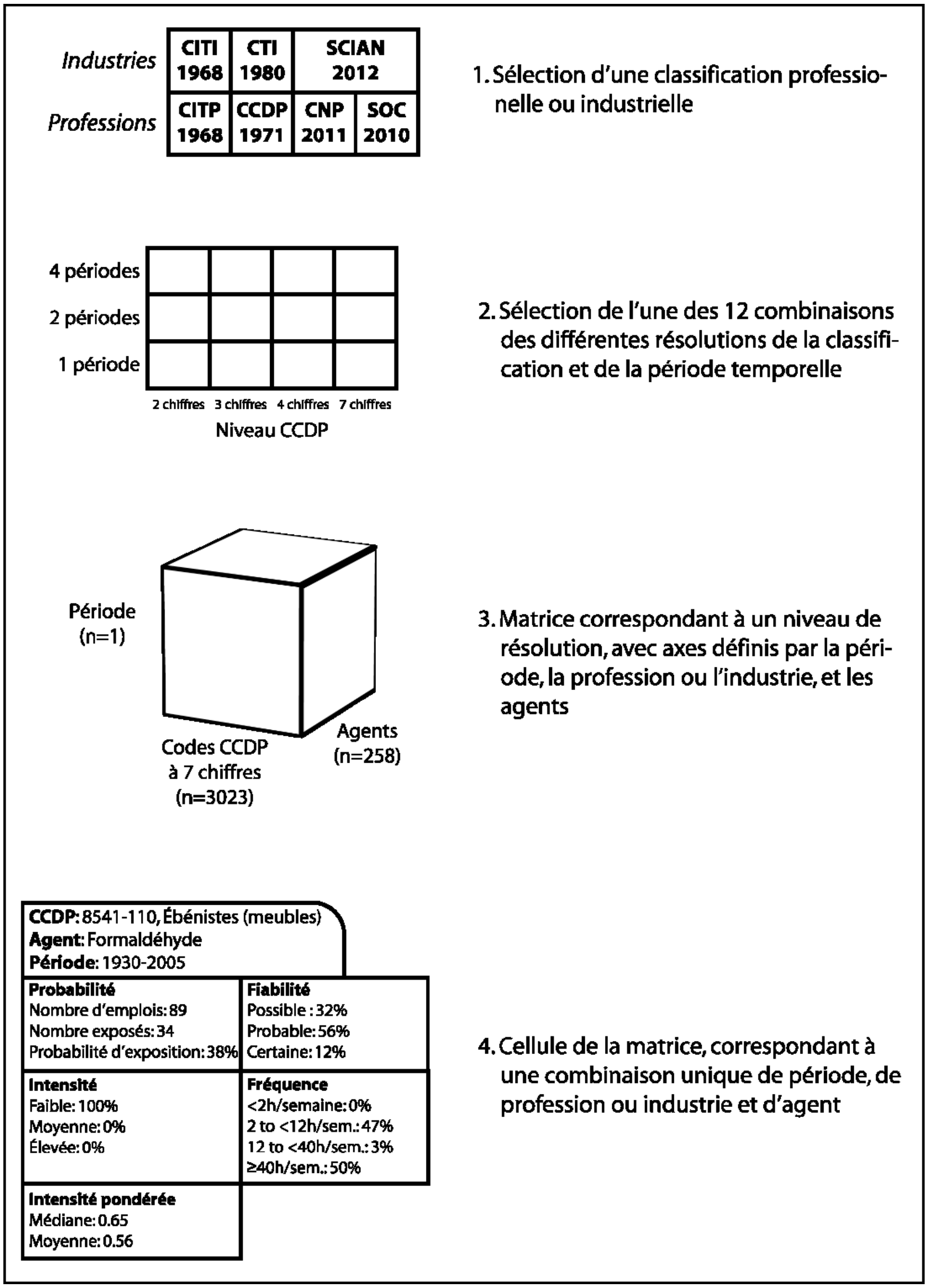
2.2 Résumé de la méthodologie par chapitre

2.2.1 Développement de la matrice CANJEM

L'objectif du projet CANJEM est de rendre disponible l'information sur les évaluations par expertise assignées au cours des études montréalaises sous forme de MEE afin d'en faciliter l'application par des utilisateurs externes. Les étapes de préparation des données ont inclus le codage de chacun des 31 673 emplois en des classifications additionnelles, portant le total à quatre classifications canadiennes et internationales pour les professions et trois pour les industries (Annexe 2). Une liste de 258 agents (présentés à l'Annexe 3) évalués communément dans les quatre études a été définie, et un indice continu combinant l'intensité et la fréquence d'exposition normalisée sur une semaine de travail de 40 heures, soit l'intensité d'exposition moyenne pondérée, a été calculé. Afin de permettre le calcul de cet indice, des poids numériques ont dû être attribués aux catégories d'intensité faible, moyenne et élevée. Différentes échelles pondérales ont été explorées pour estimer les ratios entre les catégories, allant d'une échelle linéaire (1 : faible, 2 : moyen et 3 : élevé) donnant une plus grande influence à la fréquence d'exposition, à des échelles exponentielles (1-3-9, 1-5-25 et 1-10-100) donnant une plus grande influence à l'intensité dans le calcul de cet indice. Bien que ces poids relatifs puissent varier d'un agent à l'autre dans les études du groupe, l'échelle 1-5-25 a été jugée comme étant la meilleure représentation globale des ratios entre les catégories au travers des agents après consultation avec les experts. Par ailleurs, une évaluation réalisée dans le cadre des travaux menés au Chapitre 3 a montré des corrélations de Kendall élevées ($\tau \geq 0.7$) dans les valeurs d'IEMP calculées à partir des différentes échelles de pondération, même entre les deux échelles extrêmes (1-2-3 et 1-10-100). Les corrélations stratifiées pour chacun des 258 agents étaient également du même ordre. Pour ces raisons, seules les valeurs d'intensité d'exposition moyenne pondérée calculées avec le facteur 1-5-25 ont été conservées pour l'élaboration de CANJEM.

Les données d'exposition ont ensuite été synthétisées pour réaliser une MEE dotée de trois axes représentés par les agents, les professions ou industries, et les périodes temporelles. L'axe des professions ou industries est défini par les sept classifications professionnelles ou industrielles, où chacune étant déclinée en de multiples niveaux de résolution allant de professions/industries précises à de groupes plus larges. L'axe temporel est également décliné en plusieurs résolutions, allant d'une seule période (1930-2005) pouvant être stratifiée en deux périodes (1930-1969 et 1970-2005) ou quatre périodes (1930-1949, 1950-1969, 1970-1984 et 1985-2005). Une cellule de CANJEM représente une combinaison unique de ces trois axes. La Figure 2 présente une schématisation de l'organisation des différents niveaux d'information à l'intérieur de CANJEM allant de la sélection d'un système de classification jusqu'à une cellule spécifique.

Figure 2. Organisation de la matrice CANJEM, et illustration du processus de sélection d'une cellule basée sur la classification CCDP



Chaque cellule de CANJEM comporte cinq indices qui décrivent le profil d'exposition des emplois évalués au cours des études. La probabilité représente le pourcentage d'emplois exposés à un agent relativement au nombre total d'emplois de la cellule. Les indices de fiabilité, intensité et fréquence sont quant à eux représentés par des pourcentages relatifs des emplois exposés à travers les catégories. Pour la fréquence, le nombre d'heures d'exposition hebdomadaire, exprimé sous forme continue, a été catégorisé en quatre niveaux, soit moins de 2 heures, 2 à 12 heures, plus de 12 à moins de 40 heures et 40 heures par semaine ou plus. Finalement, l'indice continu de l'intensité d'exposition moyenne pondérée est représenté par les valeurs médiane et moyenne des emplois exposés.

2.2.2 Modélisation des évaluations d'experts

L'utilisation des informations de CANJEM à un niveau de résolution précis peut faire en sorte que les estimations de l'exposition d'une cellule soient associées à une incertitude élevée lorsqu'elle est basée sur un petit nombre d'emplois. Les travaux présentés au Chapitre 4 présentent une approche visant à augmenter la précision des estimations en utilisant l'information contenue dans les cellules de professions similaires, et en organisant ce partage d'information entre les cellules selon la structure hiérarchique de la classification CCDP. Ces travaux ont porté sur les indices de la probabilité d'exposition et de l'intensité d'exposition moyenne pondérée des cellules à l'aide de modèles de régression logistique et de régression linéaire, respectivement.

Le Tableau II illustre la distribution des moyennes géométriques de l'intensité d'exposition moyenne pondérée pour le formaldéhyde durant la période 1970-1984 pour l'ensemble des groupes nichés dans le sous-groupe des travailleurs spécialisés dans la fabrication, le montage et la réparation d'articles en bois. Pour une cellule basée sur un faible nombre de données, telle la profession de Monteur-ébéniste (n=2), l'approche visait à utiliser l'information contenue dans les autres cellules du groupe hiérarchique, soit celui des Ébénistes et menuisiers en meubles, pour obtenir une estimation plus précise.

Des modèles hiérarchiques bayésiens, comportant les niveaux des sous-groupes, groupes de base et professions, ont été appliqués pour agréger les estimations de la probabilité d'exposition des cellules ou des niveaux d'intensité d'exposition moyenne pondérée des emplois exposés (après transformation logarithmique). L'estimation ou la prédiction pour un indice d'exposition pour chaque cellule pouvait ainsi être réalisée simultanément à travers ces trois niveaux avec un même modèle. Les modèles ont par ailleurs été appliqués séparément aux données pour chacune des 4 périodes constituant l'axe temporel le plus précis de CANJEM.

L'utilisation de la modélisation a également permis d'explorer l'influence de l'étude à la source des évaluations à l'aide d'une variable binaire comprenant l'étude multisite comme première catégorie, et une autre catégorie regroupant les trois autres études réalisées quelques 15 ans plus tard et axées sur un site de cancer spécifique.

Tableau II. Nombre d'emplois exposés au formaldéhyde et moyennes géométriques de l'intensité d'exposition moyenne pondérée pour les cellules nichées à l'intérieur du sous-groupe CCDP 854 (Travailleurs spécialisés dans la fabrication, le montage et la réparation d'articles en bois), période 1970-1984

Code CCDP	Description	N[1]	IEMP (MG)[2]
854	**Travailleurs spécialisés dans la fabrication, le montage et la réparation d'articles en bois**	**33**	**0,41**
8540	**Contremaîtres de travailleurs spécialisés dans la fabrication, le montage et la réparation d'articles en bois**	**7**	**0,40**
8540-110	Contremaître d'ébénistes et de menuisiers en meubles (meubles; travail du bois)	6	0,33
8540-114	Contremaître de contrôleurs, vérificateurs et trieurs de la fabrication, du montage et de la réparation d'articles en bois (meubles)	1	1,25
8541	**Ébénistes et menuisiers en meubles**	**25**	**0,44**
8541-110	Ébéniste (meubles)	17	0,31
8541-126	Réparateur de menuiseries d'assemblage (meubles; travail du bois)	1	1,00
8541-138	Encolleur (meubles; travail du bois)	1	1,00
8541-150	Monteur de meubles (meubles)	2	1,00
8541-156	Monteur-ébéniste (meubles)	2	0,35
8541-178	Ouvrier à la presse à laminer (meubles; travail du bois)	1	5,00
8541-210	Monteur d'articles en bois (meubles; travail du bois)	1	1,00
8546	**Contrôleurs, vérificateurs et trieurs de la fabrication, du montage et de la réparation d'articles en bois**	**1**	**0,13**
8546-199	Autres contrôleurs, vérificateurs et trieurs de la fabrication, du montage et de la réparation d'articles en bois	1	0,13

1. Nombre d'emplois exposés au formaldéhyde par cellule

2. Moyenne géométrique des valeurs d'intensité d'exposition moyenne pondérée des emplois exposés au formaldéhyde

2.2.3 Estimation de niveaux quantitatifs d'intensité d'exposition aux poussières de bois dans CANJEM

Le Chapitre 5 présente une approche permettant d'attribuer des niveaux quantitatifs pour l'intensité de l'exposition des cellules de CANJEM en utilisant l'information contenue dans la banque CWED. Parmi les agents communs aux deux sources, les poussières de bois représentaient un point de départ intéressant en raison de sa prévalence d'exposition relativement élevée à l'échelle nationale et provinciale (Labrèche et coll., 2012; Peters et coll., 2015), de la disponibilité de mesures dans CWED (n=6569), et des méthodes d'échantillonnage et d'analyse relativement simples, basées sur la gravimétrie (ou mesure pondérale) (Drolet et Beauchamp, 2012).

Les mesures de poussières de bois de la banque CWED ont notamment été restreintes à celles dont le titre d'emploi (selon la classification CNP 2006) était disponible. Les codes CNP 2006 ont ensuite été convertis vers une version plus récente de cette classification (2011) afin de permettre l'arrimage avec CANJEM, dont la version utilisée était définie par le niveau des groupes de base de la classification CNP 2011 (codes à 4 chiffres) et la période 1930-2005. L'analyse conjointe de ces deux sources a porté sur l'information correspondante à 31 groupes de base, lesquels au moins 10 mesures étaient disponibles dans CWED et au moins 1 emploi était exposé dans la cellule de CANJEM. Un total de 5170 mesures ont été conservées, couvrant la période entre 1978 et 2001.

L'estimation des niveaux d'exposition par profession et par catégorie d'intensité des cellules de CANJEM a été basée sur l'approche présentée par Friesen et coll. (2012) et Peters et coll. (2011b). Elle consiste à ajuster un modèle hiérarchique aux concentrations, comprenant à la base une variable pour la catégorie d'intensité des cellules d'une MEE en tant qu'effets fixes, et les professions et/ou industries entrées comme effets aléatoires. Ce modèle a été adapté dans ce travail pour tenir compte de la distribution relative des emplois entre les trois catégories d'intensité dans les cellules de CANJEM, en utilisant une variable pour la proportion d'emplois exposés à intensité moyenne, et une autre variable pour la proportion d'emplois exposés à intensité élevée. Le modèle a également inclus la durée, l'année et la zone de mesure (respiratoire ou ambiante) de l'exposition, ainsi de que la banque de données provinciale à l'origine des mesures. Finalement, les paramètres estimés du modèle ont permis de prédire des moyennes géométriques pour les concentrations aux poussières de bois pour chacune des 31 professions représentées par des mesures, pour une exposition sur 8 heures en zone respiratoire pour l'année 1989 et en tenant compte des niveaux d'intensité dans la cellule. Des prédictions ont également été réalisées pour l'ensemble des professions avec une probabilité d'exposition non nulle dans CANJEM à partir de la distribution relative des niveaux d'intensité calibrés.

2.2.4 Comparaison des expositions assignées avec l'approche hybride PROtEuS à celles assignées par la méthode par expertise traditionnelle

L'approche par expertise hybride développée dans le cadre de l'étude PROtEuS visait à accroître l'efficacité et la cohérence de l'évaluation par expertise en utilisant une synthèse d'évaluations passées comme source de référence. Elle pouvait en outre réduire les chances qu'une exposition particulière soit oubliée par l'expert. Toutefois, puisque des encodages antérieurs compilés dans les profils étaient proposés systématiquement à titre de guide, il est possible que les experts aient eu tendance à les appliquer sans tenir compte suffisamment des informations individuelles recueillies lors des entrevues, ce qui atténuerait la variabilité dans les expositions assignées aux emplois à l'intérieur d'une profession donnée, de manière semblable à l'application d'une MEE.

Afin d'évaluer l'impact de l'application de l'approche hybride sur le nombre, la fiabilité et l'hétérogénéité des expositions assignées par les experts, une comparaison des expositions assignées dans l'étude PROtEuS a été menée en utilisant les évaluations des emplois de l'étude poumon (restreintes aux sujets masculins) comme groupe de référence. La comparaison a porté sur les données d'exposition à 203 agents pour les emplois associés à 90 professions de cols bleus. Les comparaisons ont été menées d'une part sur la base des emplois et des expositions individuelles, et d'autre part, sur une agrégation des expositions par combinaison de professions et d'agents.

Les comparaisons basées sur les données individuelles visaient à évaluer si les emplois évalués par l'approche hybride étaient exposés à un plus grand nombre d'agents différents comparativement à l'approche par expertise traditionnelle, et avec un niveau de fiabilité accru. Ce dernier aspect a été évalué par des modèles de régression ordinale ajustés sur les catégories de fiabilité des expositions, en utilisant l'étude comme variable explicative. Ce modèle a également été utilisé pour comparer les niveaux d'intensité et de fréquence d'exposition des emplois entre les deux approches d'évaluation. Une autre série de comparaisons a été menée en utilisant les expositions agrégées par combinaison de professions et d'agents comme unité d'analyse (p.ex. plombiers :plomb), afin d'évaluer les différences dans la variabilité des niveaux assignés par indice d'exposition.

2.3 Modélisation statistique des données d’exposition

Les travaux présentés dans les chapitres suivants ont fait appel à une variété d’approches de modélisation statistique. La présente section résume les principaux types de modèles utilisés (linéaires, logistiques et ordinaux) et présente une brève introduction aux modèles hiérarchiques. Finalement, les principes de l’analyse bayésienne font l’objet d’une description en lien à leur application dans ce travail.

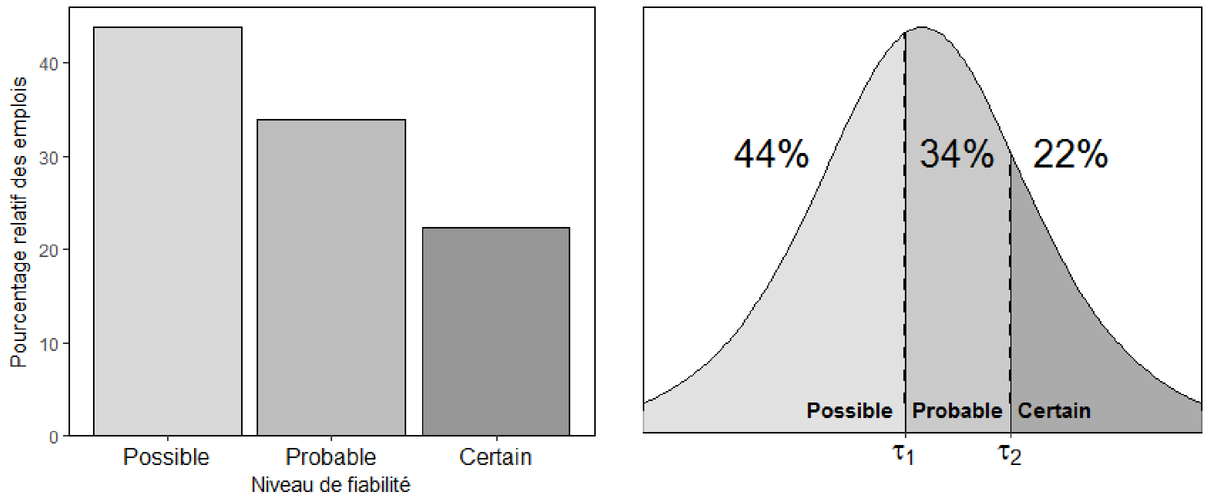
2.3.1 Modèles de régression linéaire, logistique et catégorielle

Les modèles de régression linéaire sont une approche statistique visant à expliquer l’association entre une variable dépendante et une ou plusieurs variables indépendantes (ou prédictives). Ces modèles permettent d’une part d’estimer l’influence de chaque variable sur la réponse, et de réaliser des prédictions sur la réponse à partir de scénarios basés sur des combinaisons de variables. Cette approche a été utilisée au Chapitre 5 pour estimer des niveaux d’exposition quantitatifs aux catégories d’intensité des cellules de CANJEM en ajustant un modèle aux concentrations en poussières de bois de la banque CWED. La liste de variables prédictives du modèle incluait la durée et l’année de la mesure, en plus de la distribution des catégories d’intensité dans la cellule par profession. L’estimation de l’influence relative de ces paramètres sur les niveaux d’exposition a ensuite permis de prédire des concentrations moyennes sur 8 heures pour une année et une profession donnée. Pour les analyses réalisées au Chapitre 4, les variables du modèle ajusté aux valeurs d’intensité d'exposition moyenne pondérée des emplois comprenaient les sous-groupes, groupes de base et professions de la classification CCDP, et l’étude, en utilisant une approche hiérarchique décrite plus en détails à la section suivante. Par ailleurs, tant les concentrations en poussières et les valeurs d’intensité d'exposition moyenne pondérée ont fait l'objet d'une transformation logarithmique avant l'ajustement du modèle afin d'obtenir une variable réponse distribuée de manière approximativement normale.

Les travaux ont également impliqué l'utilisation de modèles logistiques, adaptés à une variable réponse de nature binaire, tel le statut de malade/non-malade ou le statut exposé/non-exposé d'un emploi. Cette approche a été utilisée au Chapitre 4 pour modéliser la probabilité d'exposition des cellules de CANJEM, pour lesquelles les valeurs possibles sont restreintes à l'intervalle 0-100%. Les modèles logistiques ont également été appliqués dans l'article présenté à l'Annexe 1 pour estimer le risque de cancer de la prostate associé à l'emploi dans une profession ou industrie donnée. L'association entre une variable prédictive et la réponse dans les modèles logistiques est interprétée par un rapport de cotes. En prenant comme exemple l'analyse du risque de cancer de la prostate par profession ou industrie, le rapport de cotes pour une profession donnée représente la cote de maladie chez les sujets exposés (c.-à-d. ceux ayant occupé un emploi dans la profession durant au moins un an au cours de la carrière), divisé par la cote de maladie chez les non-exposés.

Finalement, des modèles adaptés à une variable réponse de nature ordinale (modèles à cotes proportionnelles) ont été utilisés pour comparer la distribution relative des catégories de fiabilité, d'intensité et de fréquence d'exposition assignées aux emplois entre l'étude PROtEuS et l'étude poumon. La variable réponse peut être vue comme une catégorisation d'une distribution continue sous-jacente, tel qu'illustré à la Figure 3, à l'aide d'une distribution logistique, délimitée par des points de coupures (τ) entre les catégories.

Figure 3. Représentation d'une distribution latente continue associée à une distribution de pourcentages relatifs par catégorie



Les modèles à cotes proportionnelles permettent d'estimer des associations sous forme de rapports de cotes cumulatifs. L'évaluation de la différence dans la distribution relative des catégories de fiabilité a été réalisée en modélisant les scores assignés aux emplois exposés par une variable binaire pour l'étude, soit l'étude poumon comme modalité de référence, ou l'étude PROtEuS comme autre modalité. La direction générale des associations suit l'interprétation des rapports de cotes estimés par un modèle logistique : un rapport de cotes cumulatif supérieur à 1 traduit ainsi une proportion relative plus élevée d'expositions assignées avec une fiabilité plus grande dans PROtEuS, relativement aux expositions assignées dans l'étude poumon.

2.3.2 Modèles hiérarchiques

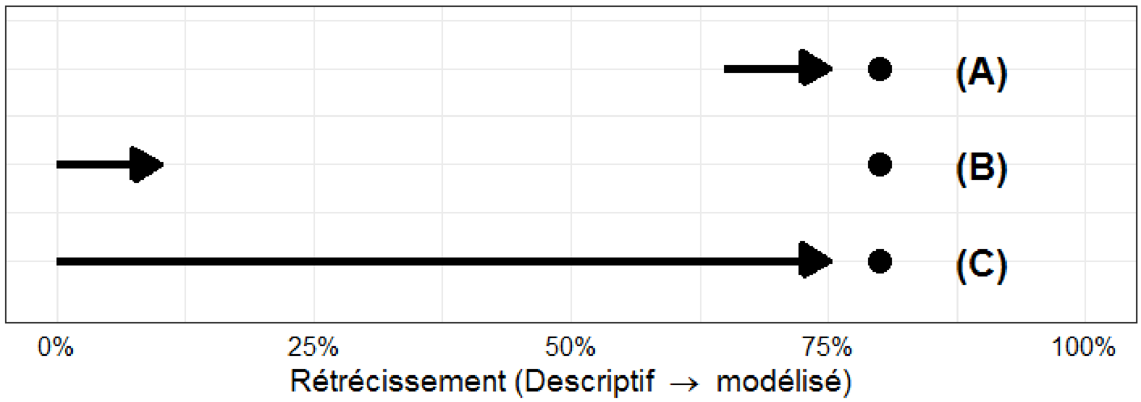
Certaines analyses ont mis en application des modèles hiérarchiques ou modèles multiniveaux. Les modèles hiérarchiques sont principalement utilisés pour tenir compte de la corrélation ou le regroupement des données provenant d'une même unité. Par exemple, ces modèles ont été utilisés lors de l'exploitation d'une banque contenant plus de 20 000 données de concentration en créatinine et de densité spécifique mesurées par le laboratoire de l'IRSST pour tenir compte de la corrélation entre les mesures prises chez un même travailleur, dans le cadre de travaux connexes menés durant la réalisation de cette thèse (Sauvé et coll., 2015). Les données peuvent également être regroupées sur la base d'une même région (Peters et coll., 2011b; Peters et coll., 2016) ou d'une même étude, dans le cas de méta-analyses (Lavoué et coll., 2007; Olsson et coll., 2011). Ces modèles permettent de considérer les coefficients des catégories d'une variable (p.ex. catégorie identifiant un travailleur) comme venant d'une certaine distribution globale plutôt que représentant chacune un ensemble fini de valeurs indépendantes. Comparativement à une approche non-hiérarchique, où l'estimation pour chaque catégorie est indépendante des autres, les coefficients des catégories entrées dans un modèle hiérarchique sont tirés à différents degrés vers la moyenne globale de la distribution, un effet appelé « rétrécissement » (« shrinkage »).

Les modèles hiérarchiques ont été appliqués dans le Chapitre 4 pour tirer profit de ce phénomène. À titre illustratif, la moyenne géométrique de l'intensité d'exposition moyenne pondérée pour la profession de monteur d'articles en bois au Tableau I est associée à une faible précision puisqu'elle est basée sur un seul emploi. Afin d'obtenir une estimation plus précise, la moyenne pour cette profession peut être conçue comme venant d'une distribution plus large englobant les autres professions du groupe des ébénistes et menuisiers en meuble. Le partage d'information sur l'exposition entre ces professions peut donc permettre d'estimer une moyenne géométrique plus précise pour les monteurs d'articles en bois, qui sera tirée (par le phénomène de rétrécissement) vers la moyenne du groupe des ébénistes et menuisiers en meubles. Ce processus de partage d'information est fondé sur la notion de possibilité d'échange (« exchangeability ») (Greenland, 2000) : en l'absence d'information permettant de départager les professions individuelles, la meilleure estimation possible est celle associée à la moyenne du groupe hiérarchique supérieur. À l'inverse, la disponibilité d'un grand nombre d'emplois pour une profession permet d'estimer une moyenne relativement précise, même sans considérer l'information disponible dans les autres professions appartenant au même groupe.

L'utilisation de modèles hiérarchiques pour estimer des niveaux d'exposition moyens par profession est principalement présentée dans la littérature comme une approche permettant de faciliter l'inclusion de catégories avec une faible taille d'échantillon. Selon une approche de modélisation non-hiérarchique, leur inclusion mènerait à des estimations instables (Friesen et coll., 2006) et/ou à des problèmes d'ajustement de modèle. L'approche hiérarchique permet une estimation plus précise pour les catégories de ce type, contrebalancée par le phénomène de rétrécissement vers la moyenne du groupe hiérarchique supérieur. L'amplitude pratique du phénomène de rétrécissement demeure toutefois peu documentée dans la littérature. Or, un effet de rétrécissement très fort pour une profession donnée implique que son estimation provient essentiellement des autres professions, où le gain sur la précision est contrebalancé par un biais important.

La Figure 4 présente une illustration théorique du phénomène de rétrécissement pour trois scénarios différents en prenant pour exemple la probabilité d'exposition. Les valeurs obtenues selon une approche descriptive sont représentées par le début de la flèche, et celles estimées par le modèle sont représentées par la pointe de la flèche. Les points représentent quant à eux la moyenne du groupe hiérarchique supérieur, et donc la direction du rétrécissement.

Figure 4. Illustration de trois scénarios pour le phénomène de rétrécissement, appliqué à l'indice de la probabilité d'exposition



Le scénario A est associé à un rétrécissement relativement mineur, et peut s'expliquer par la différence relativement faible entre la valeur descriptive et celle du groupe supérieur. Pour le scénario B, l'ampleur du rétrécissement est identique à celui du scénario A, mais l'estimation est très différente de la moyenne du groupe. Le faible rétrécissement représenté peut être dû à une taille d'échantillon relativement élevée. Finalement, le scénario C représente un cas où l'ampleur du rétrécissement est beaucoup plus grand, où l'estimation est fortement tirée vers la moyenne du groupe. Un effet de rétrécissement fort illustré par le scénario C pour une profession donnée pourrait être interprété de deux manières différentes. D'un côté, l'exposition beaucoup plus élevée (ou beaucoup plus faible) des quelques emplois évalués pour cette profession pourraient représenter des circonstances plus rares, et le rétrécissement vers la moyenne du groupe permet d'obtenir une estimation plus fiable. D'un autre côté, le rétrécissement fort pourrait suggérer que l'exposition de la profession et l'exposition du groupe ne sont pas similaires, et leur regroupement sur la base de la classification professionnelle ne reflète pas correctement les similitudes dans l'exposition. Puisque l'évaluation de ces deux interprétations n'était pas envisageable pour chacune des cellules de CANJEM, une partie des travaux a consisté à évaluer ce phénomène de rétrécissement pour identifier un niveau de compromis dans la taille d'échantillon des cellules permettant une influence modérée, et non extrême (tel celui représenté par le scénario C), de ce phénomène dans l'estimation des niveaux d'exposition.

Les modèles hiérarchiques ont également été appliqués dans l'estimation de niveaux quantitatifs d'exposition aux poussières de bois aux cellules de CANJEM par les mesures de la banque CWED au Chapitre 5, où les professions étaient entrées dans le modèle en tant qu'effet aléatoire. Cette approche permettait d'augmenter la précision des concentrations moyennes des professions associées à un nombre relativement faible de mesures en donnant une influence plus grande aux catégories d'intensité de la cellule dans CANJEM. En contrepartie, l'influence des niveaux d'intensité des cellules sur les moyennes géométriques prédites était moindre pour les professions disposant d'un nombre plus élevé de mesures.

Finalement, l'effet de rétrécissement associé aux modèles hiérarchiques a été utilisé pour tenir compte de la problématique des comparaisons multiples dans l'évaluation du risque de cancer de la prostate par profession et secteur industrie (article additionnel, présenté à Annexe 1). Ces ajustements ont été réalisés en raison du grand nombre de professions et de secteurs industriels évalués dans ces analyses, qui peuvent entraîner plusieurs associations faussement positives dues à des fluctuations aléatoires d'échantillonnage. L'approche retenue, de type « semi-Bayes » (Steenland et coll., 2000; Momoli et coll., 2010), consiste à considérer les rapports de cotes observés comme provenant d'une même population, définie par une distribution normale dont la variance est spécifiée *a priori*. En l'absence d'évidences fortes dans la littérature entre le risque de cancer de la prostate et diverses circonstances professionnelles, une variance de 0.25 sur le logarithme des rapports de cotes a été définie, représentant un intervalle à 95% situé entre

0.38 et 2.66. Cet ajustement avait pour effet de tirer les estimations plus incertaines vers la moyenne globale, permettant ainsi d'identifier les associations les plus robustes pouvant servir de pistes de recherche sur les facteurs de risques associés.

2.3.3 Méthodes bayésiennes

L'inférence bayésienne représente une approche probabiliste permettant d'intégrer les connaissances a priori sur la distribution d'un paramètre dans les inférences. Cette approche est utilisée entre autres en hygiène industrielle pour allier le jugement de l'hygiéniste quant à l'acceptabilité de l'exposition en milieu de travail à des mesures quantitatives, permettant ainsi de faciliter la prise de décision à partir d'un nombre limité de données (Hewett et coll., 2006; Ramachandran, 2008; Vadali et coll., 2009). L'approche bayésienne a également été utilisée dans l'évaluation de l'exposition rétrospective en alliant l'expertise d'hygiénistes à des mesures historiques de l'environnement de travail (Ramachandran et Vincent, 1999; Ramachandran, 2001).

Le principe de l'inférence bayésienne tire ses origines du théorème de Bayes, basé sur des probabilités conditionnelles, défini ci-dessous.

$$p(\theta|Y) = \frac{p(Y|\theta) \times p(\theta)}{p(Y)} \qquad \textbf{(1)}$$

Où $p(\theta)$ représente la distribution *a priori* sur la valeur d'un paramètre d'intérêt theta, $p(Y|\theta)$ représente la distribution des observations conditionnelles au paramètre d'intérêt (ou vraisemblance). La notation $p(\theta|Y)$ représente la distribution *a posteriori* du paramètre d'intérêt, conditionnelle aux observations et à la distribution *a priori*. $p(Y)$ représente la probabilité marginale de Y, également connue sur le vocable « constante de normalisation », permettant d'obtenir une aire sous la courbe de 1 pour la distribution de la probabilité *a posteriori*.

La distribution *a posteriori* combine ainsi l'information contenue à la fois dans les observations récoltées et dans la connaissance ou jugement *à priori* sur la distribution du paramètre d'intérêt, extérieur aux données. Une distribution *a priori* plus diffuse, ou moins informative, donne un poids plus grand aux données empiriques sur la distribution *a posteriori,* et inversement pour une distribution *a priori* plus informative combinée à une faible taille d'échantillon. L'estimation de la distribution *a posteriori* sur un ou plusieurs paramètres dans des analyses multidimensionnelles requiert généralement l'emploi de logiciels spécialisés tels WinBUGS (*Bayesian inference Using Gibbs Sampling* pour Windows) (Lunn et coll., 2000) et JAGS (*Just Another Gibbs Sampler*) (Plummer, 2003). Ces logiciels utilisent des approches de simulation Monte-Carlo à chaînes de Markov appliquées aux distributions conditionnelles des paramètres afin de définir la distribution *a posteriori*.

Les modèles utilisés au Chapitre 5, permettant d'estimer des niveaux quantitatifs d'exposition aux poussières de bois pour les cellules de CANJEM à l'aide des mesures de la banque CWED, ont été appliqués dans un cadre bayésien, en raison principalement de la grande flexibilité permise par les méthodes probabilistes de simulation Monte-Carlo. Les modèles ont notamment permis de tester l'application d'une contrainte pour associer des concentrations plus élevées associées à une augmentation de la proportion d'emplois exposés à des niveaux d'intensité moyenne et élevée, comparativement à un niveau d'intensité faible, par la définition des distributions *a priori* sur les coefficients. Le traitement des concentrations sous la limite de détection a également été réalisé selon une méthode d'imputation multiple à l'intérieur du modèle (Plummer, 2003; Huynh et coll., 2016) qui représente une méthode plus valide que les approches par substitution (Helsel, 2010). Finalement, les modèles hiérarchiques permettant le partage d'information sur l'exposition entre les cellules de CANJEM au Chapitre 4 ont aussi été appliqués sous un cadre bayésien. La plus grande souplesse associée aux méthodes bayésiennes a facilité le développement et l'application de modèles dotés d'une structure hiérarchique complexe à des groupes comportant régulièrement une très faible taille d'échantillon, moins compatibles avec une approche fréquentiste.

Chapitre 3. CANJEM: a general population job exposure matrix based on past expert assessments of exposure to over 250 agents

CANJEM: a general population job exposure matrix based on past expert assessments of exposure to over 250 agents

Jean-François Sauvé [1,2], France Labrèche [1,2,3], Lesley Richardson [2], Michel Gérin [1], Aude Lacourt [4], Tracy L Kirkham [5], Mark S Goldberg [6,7], Marie-Claude Rousseau [2,3,8], Marie-Élise Parent [2,3,8], Jack Siemiatycki [2,3], Jérôme Lavoué [1,2]

1. Université de Montréal, School of Public Health, Department of environmental and occupational health, Montréal, Québec, Canada
2. Centre de recherche du CHUM, Montréal, Québec, Canada
3. Université de Montréal, School of Public Health, Department of Social and Preventive Medicine, Montréal, Québec, Canada
4. Université de Bordeaux,Institut de santé publique d'épidémiologie et de développement (ISPED), Bordeaux, France
5. University of Toronto, Dalla Lana School of Publc Health, Toronto, Ontario, Canada
6. McGill University, Department of Medicine, Montréal, Québec, Canada
7. McGill University Health Centre, Division of Clinical Epidemiology, Montréal, Québec, Canada
8. INRS-Institut Armand-Frappier, Université du Québec, Laval, Québec, Canada

Jean-François Sauvé a contribué de façon majeure à l'analyse des données et à l'interprétation des résultats, ainsi qu'à la rédaction et à la préparation du manuscrit.

3.1 Abstract

Objectives: Information on retrospective occupational exposure covering a wide range of substances and industries is limited. We developed a job-exposure matrix (JEM) using data available in the form of expert evaluations from four population-based case-control studies of cancer (lung, breast, brain, multisite) conducted in Montreal since the 1980s.

Methods: CANJEM summarizes exposure information from 31,673 jobs held between 1930-2005. For each job, experts had evaluated the intensity, frequency and likelihood of exposure to a predefined list of agents based on jobs histories and descriptions of tasks and workplaces. CANJEM is defined by three dimensions: agents (n=258), occupation or industry (7 Canadian and international classifications, each in several resolutions) and period (1, 2 or 4 categories). Each cell provides an estimated probability of exposure (P) and summaries of the likelihood, frequency, intensity and frequency-weighted intensity (FWI) among exposed jobs.

Results: Using CANJEM defined by 4-digit occupations and period 1930-2005, the proportion of the Canadian working population covered by the cells was 91% in 1986 and 90% in 2011. Some of the most frequently-encountered agents include Polycyclic aromatic hydrocarbons (24% of jobs being exposed), organic solvents (18%), lead (13%) and formaldehyde (11%). These also had a large proportion of cells with P≥5%, while Nitroglycerine, coke dust and RDX (used primarily as an explosive) had very few cells with P≥5% but had the largest FWI values.

Conclusions: CANJEM represents one of the largest sources of retrospective exposure information currently available in terms of agents and period covered, and is available freely on the web at www.canjem.ca. Available in several classification systems, CANJEM can be used

to support exposure assessment efforts in epidemiology and prevention of occupational diseases in Canada and elsewhere.

3.2 Introduction

Assessing exposure to occupational chemical and physical agents in community-based studies needs to represent the diversity of occupations and workplaces found in the population, often over decades, while trying to achieve high validity. Due to scarcity of historical measurements, expert review of individual jobs and job-exposure matrices (JEMs) were developed to reconstruct lifetime occupational exposures in these studies (Siemiatycki *et al.*, 1981; Hoar, 1983; Gérin *et al.*, 1985; Stewart and Stewart, 1994; Siemiatycki, 1996; Teschke *et al.*, 2002). Expert review entails translating detailed job descriptions collected from interviews/questionnaires into exposure estimates that accounts for individual characteristics in tasks and other occupational factors. As this approach is expensive and time-consuming, JEMs may represent a more economical alternative. A JEM is organized as a cross-tabulation of standardized occupation and/or industry titles linked with exposures, where each cell provides estimates of exposure to one or several agents. A JEM can be applied to any study population for which occupation/industry titles have been collected. The main theoretical drawback of a JEM is that it fails to account for potential heterogeneity in exposure profiles between individual jobs within an occupational group.

Very few multi-occupation, multi-agent generic JEMs are currently in use. Notable examples include the French MATGÉNÉ system (Févotte *et al.*, 2011), currently containing exposure information for 17 agents, and the Finnish FINJEM (Kauppinen *et al.*, 1998; Kauppinen *et al.*, 2014), covering 74 agents (including psychosocial, physiological and ergonomics factors). FINJEM has also been adapted in other countries (Kauppinen *et al.*, 2009; García *et al.*, 2013; van Tongeren *et al.*, 2013).

Since the 1980s, our group has been involved in conducting four large, population-based case-control studies in the Montreal metropolitan area and other Canadian cities. In each study, expert review was used to assess exposure to approximately 300 different agents in various physical forms according to the information collected from subjects using interviews and specialized questionnaires covering each job held over their working life (Gérin *et al.*, 1985). These studies included altogether more than 8,000 subjects representing over 30,000 jobs held since the 1930s. This database represents a unique source of information about occupational exposures in a mostly urban North American population in the late 20th century, that would be very valuable as a shared source of information for researchers and others interested in assessing occupational exposures. To facilitate the dissemination and use of this data, we organized the individual job information from our studies into a JEM that we call The *Canadian Job Exposure Matrix*, or CANJEM.

This paper describes the development of CANJEM, starting with the pooling of exposure data from the four individual studies, to the definition of JEM dimensions and computation of exposure indices, as well as a descriptive summary of the resulting JEM.

3.3 Methods

Case-control study data

The Montreal case-control studies

The four case-control studies included in CANJEM have been described previously. Study 1 (conducted 1979-1986) investigated 19 cancer sites among men aged 35-70 years (3,726 cancer patients and 533 population controls) (Siemiatycki *et al.*, 1987). Study 2 (1996-2001) was a study of lung cancer and included males and females aged 35-75 years (1205 cases and 1541 controls) (Ramanakumar *et al.*, 2007). Study 3 (1996-1997) was a study of postmenopausal breast cancer among women aged 50-75 years (608 cases and 667 controls) (Labrèche *et al.*, 2010). Study 4 (2000-2004) was a study of brain tumors, representing the Quebec and Ontario portions of the multi-centric INTEROCC study (Lacourt *et al.*, 2013), and included men and women aged 30-59 years (218 cases and 414 controls). In all studies, incident cases were actively recruited from pathology departments of hospitals in the Montréal area, while population controls were selected randomly from electoral lists (Studies 1, 2 and 4) or from women diagnosed with other cancers (Study 3) and frequency-matched to cases by age and sex.

Exposure assessment methods

The exposure assessment method developed in Study 1 is described in detail in Gérin *et al.* (1985) and Siemiatycki *et al.* (1991) and was applied in subsequent studies. Briefly, complete occupational histories including job titles, employment duration, tasks performed, work environment, products and equipment used were collected from extensive face-to-face or telephone interviews. Proxy respondents (generally spouses) provided occupational histories when subjects were unable to do so.

A team of trained experts in chemistry and industrial hygiene, unaware of the case/control status of subjects, reviewed the occupational histories to classify each job ever held according to standardized occupation and industry codes. Exposures to a predefined list of approximately 300 chemical, physical and biological agents, including mixtures and broad chemical families, were then attributed to each job. Experts split and/or combined consecutive jobs that were assumed to be relatively homogenous in exposure over time. A job was considered exposed if an agent was present in the workplace at levels above those in the general (non-occupational) environment. The experts rated exposure for each combination of job and agent according to three dimensions: reliability, intensity and frequency of exposure. Reliability, or the expert's confidence that the exposure occurred, was rated as possible, probable or definite. Intensity of exposure was rated as low, medium or high. These levels were applied on a relative scale by agent (and not explicitly defined on quantitative concentration levels), where low represented a background occupational level and high the highest levels experienced in the work environment. They were guided by benchmark occupations associated with each category as illustrated in Parent *et al.* (2007) and Vida *et al.* (2010). Lastly, frequency of exposure was rated in Study 1 using the following categories: <5%, 5-30% and ≥30% of the workweek, representing <2 hours, 2-12 hours and ≥12 hours out of a typical 40-hour workweek. In Studies 2, 3 and 4, experts attributed the number of hours per week exposed for each of the three intensity ratings. For example, a given job could have an exposure profile defined by 20 hours per week at low intensity, 20 at medium, and none at high. In all studies, each job was evaluated by two experts, and consensus was used to resolve disagreements in the exposures assigned. Periodic reviews were also conducted to ensure consistency in the assessments (Siemiatycki et coll., 1991).

Pooled exposure database

Standardized occupational and industrial classifications

In developing CANJEM, the occupation and industry coding was extended so that each of the 30,000 jobs were independently coded into the same four occupation classification and the same three industry classification systems used in Canada, North America and internationally. These classifications and their hierarchical coding structures are presented in Table I.

The coding of job and industry titles into each classification was carried out by a team of trained experts using the original job descriptions and initial codes, official documentation and a purpose-built tool available online (http://www.caps-canada.ca).

Chemical/physical agents

A total of 258 agents were coded in all four studies and included in the CANJEM database. These are shown in Supplementary Table S1, and on the CANJEM project's website at http://www.canjem.ca. The agents cover a wide range of compounds and can be specific chemicals (e.g. phosgene, styrene, ozone), mixtures (e.g., gasoline, coal dust), groups based on use (e.g. pesticides, cleaning agents), chemical classes (e.g. lead compounds, aromatic amines), or physical agents (radio and microwave, ionizing and ultraviolet radiation).

Exposure indices of individual jobs

Constructing the database involved pooling data from jobs evaluated in four studies conducted over a 25-year period. Changes in the way the intensity and frequency of exposure were expressed between studies occurred over time, thus we associated each exposed job with the following pooled indices, derived from each study specific information (Table 2): intensity (low, medium, high), reliability (possible, probable, definite), frequency (<2 hours, 2-12 hours, 12 to <40 hours, and ≥40 hours per week). Lastly, we developed frequency-weighted intensity (FWI), a continuous index that combines intensity and frequency. For each exposed job/agent pair in the database, the intensity level (using quantitative scores for low, medium and high) was multiplied by the proportion of hours exposed relative to a 40-hour workweek.

Regarding the scores applied to the low, medium and high intensity levels, our experts indicated that there were no fixed and universal guidelines to assign these categories and that the quantitative meaning of these levels varied somewhat from agent to agent. The relative quantitative levels might follow a 1: 2: 3 ratio for some agents, or a steeper trend such as 1: 10: 100 for others. It was impossible to nail down different ratios specific to each of the 258 agents, so the experts agreed that the ratio 1: 5: 25 appeared to be the best estimate of the relative meaning of low: medium: high for most situations and was retained for the computation of FWI.

Development of CANJEM

CANJEM dimensions

One CANJEM cell represents a combination of three dimensions: agent, either occupational or industrial classification, and time period (Figure 1). The agent axis includes the 258 agents described previously. For the occupation/industry dimension, CANJEM is available in one of four occupational and three industrial standardized classifications separately. For each classification, exposure estimates are provided across a range of resolutions from the most detailed categories (e.g. 5-digit codes for the 1968 International Standardized Classification of Occupations, or ISCO'68) to broader groupings (e.g. 2-digit ISCO'68 codes), as listed in Table I.

Regarding the third axis (time period), we were faced with two competing tendencies. As shorter periods were defined, the specificity and validity of the information would increase, but the number of observations in each cell would decrease. Thus to accommodate different possible levels of resolution of time periods and occupational/industrial classifications, we produced several versions of CANJEM using a single global period (1930-2005), 2 periods (1930-1969, 1970-2005) to reflect changes in the organization of occupational health and safety in Canada starting in the 1970s (Verma, 1996), and 4 time periods (1930-1949, 1950-1969, 1970-1984 and 1985-2005). CANJEM can be searched with any of those three schemes.

This organization of the occupation/industry and period axes allows the user to select for a particular situation of interest an exposure estimate from among different resolutions of the occupations/industries and time periods. This feature could be useful when no exposure information is available at the finest resolution of the JEM for a given job: one could then choose to use the estimate from a less precise occupation/industry code and/or from a broader time period (*e.g.*, from the single or two time period axis), or both.

CANJEM, rather than a single JEM, therefore represents a set of JEMs, each defined by the choice of a particular occupation or industry classification and its associated resolution, and a time period scheme (1, 2 or 4). The process of selecting a specific version of CANJEM is illustrated in figure 1.

Exposure indices of cells

Each cell in a particular version of CANJEM provides an estimated probability of exposure, and, for exposed jobs, the reliability, intensity, frequency and FWI of exposure (Table II). These indices are calculated by summarizing information from all individual jobs in the pooled database associated with the cell. A job was included in a period when the employment dates covered at least one year in the time period. Jobs with an employment period straddling two or more time periods can therefore contribute to multiple time periods.

The probability of exposure is the proportion of jobs in a given cell that were considered exposed to the agent of interest, and ranges from 0% to 100%. Exposed jobs were defined as having a frequency of exposure of at least 30 minutes per week, a reliability level of "possible" or greater, and a FWI of at least 0.05, which corresponds to 2 hours per week at low intensity.

Each cell also provides the distribution of exposed jobs (as relative percentages) across each categorical rating for reliability (possible, probable, definite), intensity (low, medium, high) and frequency (<2h, 2-<12h, 12-<40, ≥40 hours per week) of exposure. Estimates for the continuous index of FWI are provided as median and arithmetic mean values across exposed jobs in the cell.

For each JEM, all cells for which one job or more were available in the pooled exposure database are included. The selection of a specific minimum sample size per cell is to the CANJEM users' discretion.

Descriptive analyses of CANJEM

Coverage of the Montreal and Canadian populations

Since CANJEM is based on data generated from a fixed set of real subjects in our past studies, it cannot be assumed that CANJEM has exposure estimates available for every occupation or industry at any level of resolution. We therefore conducted analyses to describe the extent of coverage of CANJEM for the Montreal and Canadian populations at two different times represented in the 1986 Census of Canada (Statistics Canada, 1989) and the 2011 National Household Survey (Statistics Canada, 2016). CANJEM versions used were based on the Canadian classification specific to each census, namely the 4-digit level of the 1971 Canadian Classification and Dictionary of Occupations (CCDO) for the 1986 census, and 4-digit level of the 2011 National Occupational Classification (NOC) for the 2011 census. For illustrative purposes, the CANJEM versions used in the analysis had a minimum sample size per cell of 10, and all time period schemes were tested. The proportion of individuals employed in occupations covered by the JEMs, relative to the total number of individuals employed in each population (Canada and Montreal), was then computed.

Probability of exposure and average FWI

To present a descriptive analysis of the information contained in CANJEM, we used the 1968 International Standard Classification of Occupations (ISCO'68) classification, commonly used in occupational epidemiology (t' Mannetje and Kromhout, 2003), with 5-digit codes, a single time period and a minimum sample size per cell of 10. The analysis focused on the probability of exposure and average FWI of cells by agent.

3.4 Results

Pooled exposure database

The pooled database contained information on a total of 31,673 jobs held by 8,760 subjects between 1920 and 2005. 15,067 (47.5%) jobs were collected during Study 1, followed by Study 2 (n=10,371, 32.7%), Study 3 (n=3510, 11.1%) and Study 4 (n=2725, 8.6%). Figure 2 presents the distribution of jobs by decade stratified by study.

Of the 31,673 jobs included in the database, 22,763 (71.9%) were exposed to at least one of the 258 agents. The agent for which we identified the largest number of exposed jobs in our database was Polycyclic aromatic hydrocarbons (PAHs) from any source (n=7,651, 24.2% of all jobs). Several associated agents such as PAHs from hydrocarbons, engine emissions and carbon monoxide also had some of the largest number of exposed jobs, as listed in Table III. For 120 agents, exposure was present in fewer than 1% of jobs, 107 had 1 to <5% of jobs exposed, 18 had 5% to <10%, and 13 had 10% or more jobs exposed.

The majority (62%) of exposed job/agent combinations had a "definite" reliability level, compared to 27% for "probable" and 11% for "possible". Forty-eight percent had a frequency in the range of 2 to <12 hours per week; relative proportions for the remaining categories were 7% for <2 hours, 18% for 12 to <40 hours, and 28% for ≥40 hours per week, the latter consisting mainly of exposure 40 hours per week (87%). For intensity, more than half of the exposed job/agent combinations had low intensity (58%), compared to 34% for medium and 8% for high intensity. Table III lists the fifteen agents with the largest number of exposed jobs in the pooled database and their distribution by reliability, intensity and frequency of exposure. A listing of the full set of 258 agents accompanied by descriptive summaries of the exposure data is shown in Supplementary table S1 (available online at expostats.ca/jeanf/chapitre3/suppl_tab_s1.xls).

CANJEM

CANJEM availability

CANJEM is available in each of the 7 occupational and industrial classification systems, for all resolutions and all time periods. They can be consulted on the www.canjem.ca website on an agent by agent, occupation by occupation (or industry by industry), or cell by cell basis. The website includes a keyword search to retrieve relevant occupation/industry codes and agents, and allows users to specify various criteria (e.g. minimum number of jobs and subjects per cell, minimum reliability). Batch versions of CANJEM can also be obtained through collaborations.

Coverage of the Montreal and Canadian populations

Using data from the 1986 Census of Canada and the 1971 Canadian Classification and Dictionary of Occupations (4-digit codes) version of CANJEM, the proportions of the Montreal working population covered by JEMs defined with 1, 2 or 4 periods were 93%, 86% and 68%, respectively. For the Canadian working population, coverage for the same JEMs was slightly lower with 91%, 81% and 63%, respectively. Using the data from the most recent census (2011) and the 2011 National Occupational Classification version of CANJEM (4-digit codes), the proportion of the working population covered by the JEMs with 1, 2 and 4 time periods were 91%, 76% and 53% for the Montreal population, and 90%, 76% and 52% for the Canadian population. As an illustration of the influence of the criterion of minimum sample size per cell (set at 10 for this calculation), the previous numbers are changed to the following when choosing a minimum of 5 jobs per cell: 95%, 86% and 63% for the Montreal population, and 94%, 86% and 64% for the Canadian population.

Agents with the highest probability of exposure and average FWI

The probability of exposure was equal to or greater than 5%, a criterion used to define a particular cell as "exposed" (Kauppinen *et al.*, 1998), for 13,960 (11.6%) of CANJEM cells defined by 5-digit ISCO'68 codes with at least 10 jobs (n=467), a single time period, and 258 agents. The median probability of exposure across this subset of 13,960 cells was 13.4% (interquartile interval 7.7-30.0%, range 5-100%). Table IV presents the exposure profiles for the 15 agents with the highest proportion of exposed cells, and for the 15 agents with the highest average FWI (using the median across exposed cells). Agents with the highest proportion of exposed cells were associated with relatively low frequency and intensity of exposure. Conversely, the 15 agents with the highest average FWI values had relatively few exposed cells.

3.5 Discussion

Occupational exposure assessment is a challenging aspect of population-based studies due to the diversity of workplaces and work conditions that need to be evaluated with limited information. To address this, our group developed in the 1980s a method based on the collection of detailed job descriptions and their translation into exposure estimates to hundreds of agents by trained experts (Gérin *et al.*, 1985). This method, although providing exposure estimates specific to the intricacies of each job held by each subject, is costly (an estimated 50 expert years were used across the 4 studies) and cannot readily be applied in other investigations. In creating CANJEM, we aggregated expert evaluations accumulated over several decades into a format usable by other researchers in epidemiological and other public health investigations.

Coverage of the Montreal and Canadian populations

CANJEM was constructed from jobs held by participants enrolled in our studies. Since these represent a sample of the population, some combinations of occupations/industries and periods may not be represented in our data, as opposed to other JEMs created by assigning exposures to a list of all occupations in a population, such as FINJEM (Kauppinen *et al.*, 1998) and MATGÉNÉ (Févotte *et al.*, 2011). Nevertheless, we found very good coverage of the Canadian working population as represented in two national surveys conducted 25 years apart (1986 and 2011), with 90% or more of the working population employed in occupations included in JEMs defined by one time period for 1930-2005. As expected the proportions of occupations covered were lower when the data is split into more time periods (down to 50-60% depending on the population), and are improved by coarser resolutions of the occupation/industry classifications or less stringent sample size criteria.

Validity

CANJEM results from the aggregation of exposure estimates in a series of case-control studies held in Montreal. Its validity therefore mainly rests on the quality of the individual estimates, as well as the representativeness of the jobs in the database compared to the Montreal and Canadian working populations (or any other population one may wish to use the JEM for).

The exposures assigned by the experts have been shown to be reliable and repeatable (Goldberg et al., 1986; Siemiatycki et al., 1997). A validation trial was also conducted where our experts assessed exposure to 19 agents (12 of which are CANJEM agents, encompassing metals, solvents and hydrocarbons, among others) for 47 jobs for which some measurements were available (Fritschi et al., 2003). Between 70% and 90% of the substances known to have been present were correctly identified. In addition, the occupational histories collected by interviews and questionnaires have been found to be accurate when compared to governmental records (Baumgarten et al., 1983).

As with other sources of information on occupational exposures, CANJEM's application to any study population requires careful evaluation. The only extensive external comparison of the evaluations of the Montreal experts was conducted by Lavoué et al. (2012) between jobs from Study 2 and FINJEM for 27 agents. Prevalence and levels of exposure were often similar between the two sources for several agents such as metals or welding fumes, but disagreements were also found for agents such as flour dust and chlorinated solvents for prevalence, and toluene and benzo[a]pyrene for intensity level. Aside from differences in exposure assessment methodology, differences in true exposure conditions could also play a role in the discrepancies observed. The studies used in creating CANJEM were set in a largely urban population with a historically important textile and garment industry, and manufacturing of food and beverage products, among others (Brodeur and Galarneau, 1994). The application of CANJEM to another population should therefore account for population-specific factors in exposure.

Decisions made in designing CANJEM

The exposure information in CANJEM combines data from studies conducted at different points in time over 25 years, from jobs held by both cases and controls, as well as by men and women. Excluding data based on one or more of these factors would have resulted in fewer cells included in CANJEM, and in fewer jobs to base the exposure estimates within each cell. On the other hand, mixing information from jobs with systematic differences in exposure profiles could lead to less reliable estimates.

Concerns regarding including information from cases have been raised in the literature since differences in exposures to known risk factors for a disease and reporting of work, tasks and exposures may occur between cases and controls (Kirkham *et al.*, 2016). Using data from study 2, Kirkham *et al.* (2016) compared JEMs created from jobs held by lung cancer cases to JEMs created from population controls. The agreement between the JEMs was high for exposure status (92%-93% concordance in cells for probability ≥5%) and for the probability and intensity of exposure, suggesting that aggregating the case and control information in our study into a single JEM is justifiable given the benefits of increased sample size.

The potential differences in exposure by men and women were evaluated by Labrèche *et al.* (2015), who compared JEMs created separately from jobs held by men and jobs held by women using data from studies 2 and 3. For 91% of the 14,337 occupation-agent combinations, the probability of exposure between held by men and jobs held by women was comparable. While differences in exposure probability were observed for several agents such as engine emissions or fabric dust, most could be explained by the different distribution of jobs held by men and women across the spectrum of occupations. Within-occupation differences could often be mitigated using finer occupation/industry codes, and only a small residual proportion could be explained by different tasks done by men and women within the same occupation. Results from this evaluation did not warrant the production of sex-specific versions of CANJEM although further refinements could be made to provide estimates stratified by gender for cells where the main differences were observed.

CANJEM includes exposure data from jobs held by subjects whose occupational histories were collected from proxy respondents, which represented approximately 22% of jobs. Compared to self-respondents, exposure assigned to jobs from proxy respondents had somewhat lower reliability ratings, however the intensity, frequency and FWI values were comparable, overall not justifying excluding them from the JEM given the added sample size.

Additional methodological considerations

The pooling of exposure data from the different studies involved significant efforts in adapting some of the exposure indices and selecting compatible agents across the four studies, but differences may remain since the studies were conducted at different points in time. Most of the exposure assessment method and infrastructure was developed for Study 1, and evolved during studies 2, 3 and 4. The relative meaning of the exposure levels representative of low, medium and high intensity may have changed over time as well (Pintos *et al.*, 2012). A comparison of JEMs created from the exposure data from Studies 1 and 2 showed that exposure probability was slightly higher in Study 2 (done 10 years later) while a larger proportion of high intensity ratings were assigned in Study 1 (results not shown). We do not think that these differences warrant the use of study-specific estimates in each cell, but their evaluation and adjustment using modelling constitute an interesting development avenue for CANJEM.

Regarding the scores applied to the low, medium and high intensity levels in the computation of FWI, we also evaluated alternative ratios of 1:2:3, 1:3:9 and 1:10:100 aside from 1:5:25; pairwise Kendall correlations between FWI values computed with the different ratios for each exposed job/agent pair were very high, with the lowest correlation (tau=0.7) found between the two most extreme ratios (1: 2: 3 and 1: 10: 100). Correlations stratified by agent were similar, and did not merit the inclusion of FWI indices computed using ratios other than 1: 5: 25.

Applicability of CANJEM to population studies

As a general population tool, CANJEM can be used multiple endeavours, including worker compensation, workplace preliminary survey, estimating numbers of workers exposed (Labrèche *et al.*, 2013), evaluating burden of disease, and of course for epidemiological studies. The type and complexity of exposure metrics needed in each case might vary widely, making CANJEM's flexible dimensional design and large array of exposure indices a significant advantage. A companion paper to be submitted shortly will provide more guidance into using CANJEM in the context of epidemiological studies based on the collective experience in our group.

Conclusion

CANJEM figures among the largest sources of information on occupational exposures in North America and beyond, built from 50 expert-years of work by our team, and is accessible online.

The combination of an extensive list of agents, multiple time periods and flexible dimensioning makes it suitable for a diversity of applications in epidemiology and occupational hygiene.

Acknowledgements

Exposure assessment methods were expertly developed and implemented by Michel Gérin, Louise Nadon, Ramzan Lakhani, Denis Bégin, and Benoit Latreille. The studies would not have been possible without the able participation of a large number of research assistants and interviewers, including Marie-Claire Goulet, Jerome Asselin, and Sally Campbell. Dora Rodriguez and Elmira Aliyeva, were responsible for the work on verifying the coding of cross-walks, while Ana Gueorguieva in conjunction with Mendher Merzoug prepared the databases and analyses required in the preparation of the final matrix.

3.6 References

Baumgarten M, Siemiatycki J, Gibbs GW. (1983) Validity of Work Histories Obtained By Interview for Epidemiologic Purposes. Am J Epidemiol; 118: 583-91.

Brodeur M, Galarneau D. (1994) Three large urban areas in transition. Perspectives on Labour and Income; 6: 5.

Dominion Bureau of Statistics. (1970) Standard Industrial Classification Manual. Ottawa: Information Canada.

Févotte J, Dananché B, Delabre L *et al.* (2011) Matgene: A Program to Develop Job-Exposure Matrices in the General Population in France. Ann Occup Hyg; 55: 865-78.

Fritschi L, Nadon L, Benke G *et al.* (2003) Validation of expert assessment of occupational exposures. Am J Ind Med; 43: 519-22.

García AM, González-Galarzo MC, Kauppinen T *et al.* (2013) A job-exposure matrix for research and surveillance of occupational health and safety in Spanish workers: MatEmESp. Am J Ind Med; 56: 1226-38.

Gérin M, Siemiatycki J, Kemper H *et al.* (1985) Obtaining occupational exposure histories in epidemiologic case-control studies. J Occup Med; 27: 420-6.

Goldberg MS, Siemiatycki J, Gérin M. (1986) Inter-rater agreement in assessing occupational exposure in a case-control study. Br J Ind Med; 43: 667-76.

Hoar S. (1983) Job Exposure Matrix Methodology. Clin Toxicol; 21: 9-26.

International Labour Office (ILO). (1969) International Standard Classification of Occupations. Revised Edition 1968. Geneva: ILO.

Kauppinen T, Heikkilä P, Plato N *et al.* (2009) Construction of job-exposure matrices for the Nordic Occupational Cancer Study (NOCCA). Acta Oncol; 48: 791-800.

Kauppinen T, Toikkanen J, Pukkala E. (1998) From cross-tabulations to multipurpose exposure information systems: A new job-exposure matrix. Am J Ind Med; 33: 409-17.

Kauppinen T, Uuksulainen S, Saalo A *et al.* (2014) Use of the Finnish Information System on Occupational Exposure (FINJEM) in Epidemiologic, Surveillance, and Other Applications. Ann Occup Hyg; 58: 380-96.

Kirkham TL, Siemiatycki J, Labrèche F *et al.* (2016) Impact of aggregating exposure information from cases and controls when building a population-based job-exposure matrix from past expert evaluations. Occup Environ Med; 73: 474-81.

Labrèche F, Duguay P, Ostiguy C *et al.* (2013) Estimating occupational exposure to carcinogens in Quebec. Am J Ind Med; 56: 1040-50.

Labrèche F, Goldberg MS, Valois M-F *et al.* (2010) Postmenopausal breast cancer and occupational exposures. Occup Environ Med; 67: 263-69.

Labrèche F, Lacourt A, Lavoué J. (2015) Occupational Exposure to Chemical and Physical Contaminants: Sex-Differentiated Analysis. Montréal: Institut de recherche Robert-Sauvé en santé et en sécurité du travail.

Lacourt A, Cardis E, Pintos J *et al.* (2013) INTEROCC case--control study: lack of association between glioma tumors and occupational exposure to selected combustion products, dusts and other chemical agents. BMC Public Health; 13: 340.

Lavoué J, Pintos J, Van Tongeren M *et al.* (2012) Comparison of exposure estimates in the Finnish job-exposure matrix FINJEM with a JEM derived from expert assessments performed in Montreal. Occup Environ Med; 69: 465-71.

Parent M-É, Rousseau M-C, Boffetta P *et al.* (2007) Exposure to Diesel and Gasoline Engine Emissions and the Risk of Lung Cancer. Am J Epidemiol; 165: 53-62.

Pintos J, Parent M-É, Richardson L *et al.* (2012) Occupational exposure to diesel engine emissions and risk of lung cancer: evidence from two case–control studies in Montreal, Canada. Occup Environ Med; 69: 787-92.

Ramanakumar AV, Parent ME, Siemiatycki J. (2007) Risk of lung cancer from residential heating and cooking fuels in Montreal, Canada. Am J Epidemiol; 165: 634-42.

Siemiatycki J. (1996) Exposure assessment in community-based studies of occupational cancer. Occupational Hygiene; 3: 41-58.

Siemiatycki J, Day NE, Fabry J *et al.* (1981) Discovering Carcinogens in the Occupational Environment: A Novel Epidemiologic Approach. J Natl Cancer Inst; 66: 217-25.

Siemiatycki J, Fritschi L, Nadon L *et al.* (1997) Reliability of an expert rating procedure for retrospective assessment of occupational exposures in community-based case-control studies. Am J Ind Med; 31: 280-86.

Siemiatycki J, Nadon L, Lakhani R *et al.* (1991) Exposure assessment. In Siemiatycki, J. Risk factors for cancer in the workplace, Boca Raton: CRC Press.

Siemiatycki J, Wacholder S, Richardson L *et al.* (1987) Discovering carcinogens in the occupational environment. Methods of data collection and analysis of a large case-referent monitoring system. Scand J Work Environ Health; 13: 486-92.

Statistics Canada. (1980) Standard Industrial Classification, 1980. Ottawa, ON: Statistics Canada, Standards Division.

Statistics Canada. (1989) Census of Canada, 1986: Occupations. Ottawa: Statistics Canada.

Statistics Canada. (2012a) National occupational classification, 2011. Ottawa: Minister of Industry.

Statistics Canada. (2012b) North American Industry Classification System (NAICS) Canada 2012. Ottawa: Statistics Canada.

Statistics Canada. (2016) 2011 National Household Survey, Statistics Canada Catalogue no. 99-012-X2011033. Ottawa, ON.

Stewart WF, Stewart PA. (1994) Occupational case-control studies: I. Collecting information on work histories and work-related exposures. Am J Ind Med; 26: 297-312.

t' Mannetje A, Kromhout H. (2003) The use of occupation and industry classifications in general population studies. Int J Epidemiol; 32: 419-28.

Teschke K, Olshan AF, Daniels JL *et al.* (2002) Occupational exposure assessment in case-control studies: opportunities for improvement. Occup Environ Med; 59: 575-94.

U.S. Bureau of Labor Statistics. (2014) Standard Occupational Classification (SOC) System. http://www.bls.gov/soc/#classification

United Nations. (1971) International standard industrial classification of all economic activities, revision 2. New York: United Nations Statistical Office.

van Tongeren M, Kincl L, Richardson L *et al.* (2013) Assessing Occupational Exposure to Chemicals in an International Epidemiological Study of Brain Tumours. Ann Occup Hyg; 57: 610-26.

Verma DK. (1996) Occupational health and safety initiatives and trends in Canada, in particular in Ontario. Ann Occup Hyg; 40: 477-85.

Vida S, Pintos J, Parent ME *et al.* (2010) Occupational exposure to silica and lung cancer: pooled analysis of two case-control studies in Montreal, Canada. Cancer Epidemiol Biomarkers Prev; 19: 1602-11.

3.7 Tables and figures

Table I. Standardized occupation and industry classifications and levels of resolution available in CANJEM

A) Occupations

Classification	Resolution	Level	Number of groups in classification
International Standardized Classification (ISCO), 1968[1,2]	1 digit	Major group	8
	2 digits	Minor group	81
	3 digits	Unit group	282
	5 digits	Occupation	1504
Canadian Classification and Dictionary of Occupations (CCDO), 1971[3]	2 digits	Major group	23
	3 digits	Minor group	81
	4 digits	Unit group	500
	7 digits	Occupation	7907
Canadian National Occupational Classification (NOC), 2011[4]	1 digit	Division	10
	2 digits	Major group	40
	3 digits	Minor group	140
	4 digits	Unit group	500
United States Standardized Occupational Classification (SOC), 2010[5]	2 digits	Major group	23
	3 digits[6]	Minor group	97
	5 digits	Broad occupation	461
	6 digits	Detailed occupation	840

B) Industries

Classification	Resolution	Level	Number of groups in classification
International Standard Industrial Classification (ISIC) revision 2, 1968[7,8]	1 digit	Major division	9
	2 digits	Division	33
	3 digits	Major group	71
	4 digits	Group	159
Canadian Standardized Industrial Classification (SIC), 1980[9]	1 digit	Division	18
	2 digits	Major group	76
	3 digits	Minor group	318
	4 digits	Unit group	860
	2 digits	Sector	20

Classification	Resolution	Level	Number of groups in classification
North American Industry Classification System (NAICS), 2012[9]	3 digits	Subsector	102
	4 digits	Group	323
	5 digits	Industry	711
	6 digits	Canadian industry	922

3. International Labour Office (ILO) (1969)
4. Includes Armed Forces as a category in each level of resolution
5. Dominion Bureau of Statistics (1970)
6. Statistics Canada (2012a)
7. U.S. Bureau of Labor Statistics (2014)
8. Level includes two 4-digit codes: 15-11 (Computer occupations) and 51-51 (Printing workers)
9. Major division 0 (Activities not Adequately Defined) and nested subgroups omitted
10. United Nations (1971)
11. Statistics Canada (1980)
12. Statistics Canada (2012b)

Table II. Exposure indices of individual jobs in the pooled exposure database and indices of the CANJEM cells

Indices in the pooled exposure database

Index	Format
Exposure status	Binary (exposed/unexposed)
Reliability [1]	Categorical (possible, probable, certain)
Intensity [1]	Categorical (low, medium, high)
Frequency [1]	Categorical (<2h, 2-<12h, 12-<40, ≥40 hours per week)
Frequency-weighted intensity [1]	Continuous

Indices in CANJEM cells

Index	Format
Probability	Percentage (proportion of jobs exposed among all jobs)
Reliability	Categorical (relative percentages of exposed jobs with possible, probable and certain reliability)
Intensity	Categorical (relative percentages of exposed jobs with low, medium and high intensity)
Frequency	Categorical (relative percentages of jobs exposed <2h, 2-12h, 12-40 and 40+ hours per week)
Frequency-weighted intensity (FWI)	Continuous (median and arithmetic average of exposed jobs)

1. Available for exposed jobs only

Table III. Number of and crude proportion of exposed jobs, proportion of exposed jobs stratified by reliability and intensity rating, and modal frequency of exposure across exposed jobs for the 15 most prevalent agents in the pooled exposure database

Agent	Number of exposed jobs (% of total)[1]	Reliability (% of jobs)[2]			Intensity (% of jobs)[2]			Frequency (hours per week)
		Possible	Probable	Certain	Low	Medium	High	Modal category (% of jobs)[2]
PAHs from any source	7651 (24%)[1]	11.8	14.2	73.9	68.8	23.7	7.5	2-<12 (34.3%)
PAHs from petroleum	5903 (19%)	4.5	14.6	80.9	69.7	23.5	6.8	12-<40 (36.3%)
Engine emissions	5816 (18%)	4.7	13.1	82.2	43.2	50.4	6.3	12-<40 (50.6%)
Organic solvents	5696 (18%)	7.2	21.2	71.7	35.1	51.5	13.4	2-<12 (55.6%)
Carbon monoxide	5298 (17%)	3.4	12.7	83.9	78.4	19.5	2.1	12-<40 (42.9%)
Lead compounds	4211 (13%)	4.7	13.7	81.5	83.5	15.4	1.1	12-<40 (49.6%)
Alkanes (C5-C17)	4056 (13%)	6.4	25.8	67.8	33.4	51.4	15.2	2-<12 (51.2%)
Aliphatic aldehydes	4047 (13%)	31.1	38.3	30.6	86.7	12.8	0.5	≥40 (51.2%)
Mononuclear aromatic hydrocarbons	3842 (12%)	6.4	21.7	72	62.3	32.8	4.9	2-<12 (47.9%)
Cleaning agents	3564 (11%)	3.5	12.4	84.1	71.7	18.1	10.2	2-<12 (69.3%)
Formaldehyde	3390 (11%)	33.4	46.0	20.6	86.4	13.0	0.6	≥40 (49.3%)
Alkanes (C18+)	3350 (11%)	6.8	23.0	70.1	49.8	33.5	16.7	≥40 (40.5%)
Metallic dust	3309 (10%)	6.3	25.0	68.7	52.5	41.0	6.5	≥40 (51.2%)
Iron	2869 (9%)	5.3	21.3	73.4	47.8	42.1	10.1	≥40 (47.1%)
Diesel engine emissions	2667 (8%)	21.4	26.4	52.2	58.8	35.5	5.7	2-<12 (42.4%)

1. Percentage of exposed jobs relative to all jobs in the CANJEM database (n=31,673)

2. Percentage of exposed jobs by category relative all exposed jobs by agent

Table IV. Fifteen agents with the largest proportion of cells with probability of exposure of 5% or greater, and highest average frequency-weighted intensity of exposure (FWI); CANJEM based on 5-digit ISCO'68 codes and period 1930-2005

Highest proportion of cells with probability ≥5% [1]							
	Probability		**Intensity**			**Frequency**	**FWI** [5]
Agent	**Cells with P≥5% (%)** [2]	**Median (%)** [3]	**Low (%)** [4]	**Medium (%)** [4]	**High (%)** [4]	**Median (h/week)**	**Mean**
PAHs from any source	**71.1**	30.5	69.1	23.7	7.3	25.0	1.00
Organic solvents	**63.4**	30.2	34.2	52.2	13.6	5.0	1.07
PAHs from petroleum	**58.7**	23.2	70.1	23.4	6.5	20.8	0.92
Alkanes (C5-C17)	**53.7**	18.8	32.4	52.2	15.4	6.0	1.32
Carbon monoxide	**53.1**	20.0	79.3	18.7	2.0	19.0	0.68
Aliphatic aldehydes	**50.7**	16.3	88.0	11.7	0.3	25.0	0.73
Mononuclear aromatic hydrocarbons	**50.5**	20.8	62.0	33.2	4.8	11.3	0.84
Engine emissions	**49.3**	15.0	41.9	51.6	6.5	9.8	0.97
Alkanes (C18+)	**48.6**	20.0	48.5	34.5	17.1	15.3	1.22
Lead compounds	**47.5**	17.7	83.8	15.3	0.9	10.2	0.57
Formaldehyde	**41.1**	15.4	28.8	87.9	11.7	0.4	0.73
Metallic dust	**37.0**	30.8	16.0	50.5	42.3	7.1	1.43
Nitrogen oxides	**35.8**	13.5	22.7	83.4	16.3	0.4	0.65
Benzo[a]pyrene	**35.8**	14.6	23.0	76.6	17.3	6.2	0.92
Iron	**35.1**	21.6	20.5	46.3	43.4	10.4	1.49
Highest average FWI , median across cells with probability ≥5%							
Agent	**Probability**		**Intensity**			**Frequency**	**FWI**
	Cells with P≥5% (%)	**Median (%)**	**Low (%)**	**Medium (%)**	**High (%)**	**Median (h/week)**	**Mean**
Nitroglycerine	0.4	5.8	0.0	14.3	85.7	40.0	**13.31**
Coke dust	1.9	10.0	30.0	25.0	45.0	45.0	**10.12**
RDX (cyclonite)	0.2	5.4	50.0	0.0	50.0	40.0	**9.52**
Coke combustion products	2.6	9.5	24.0	24.0	52.0	48.0	**5.97**
Tobacco dust	1.1	6.7	22.6	41.9	35.5	40.0	**5.65**
Fur dust	2.1	10.2	23.1	41.5	35.4	40.0	**5.47**
Trinitrotoluene	0.6	6.2	35.7	7.1	57.1	40.0	**5.00**
Sodium hydrosulphite	0.2	63.6	14.3	57.1	28.6	2.5	**4.79**
Coal tar and pitch	4.9	8.1	9.4	42.4	48.2	4.0	**4.50**
Leather dust	4.7	19.3	64.0	28.5	7.4	40.0	**3.10**
Coal dust	7.5	9.1	22.8	19.1	20.2	60.7	**2.84**
PAHs from coal	15.2	9.5	20.4	29.1	49.3	21.6	**2.69**
Coal combustion products	9.9	9.8	35.0	35.6	51.7	12.7	**2.60**
Chlorine dioxide	0.9	24.7	31.3	26.7	73.3	0.0	**2.45**

Wool fibres	8.8	35.5	40.0	60.9	35.6	3.4	**2.27**

1. CANJEM cells based on a minimum of 10 jobs (n=467)
2. Proportion of cells (out of 467) with probability of exposure ≥5%
3. Median probability across cells with probability of exposure ≥5%
4. Proportion of jobs by categorical intensity ratings across cells with probability of exposure ≥5%
5. Average frequency-weighted intensity, median value of cells with probability of exposure ≥5%

Figure 1. Illustration of the organization of the CANJEM system, using the JEM based on 5-digit ISCO'68 codes and a single time period as an example

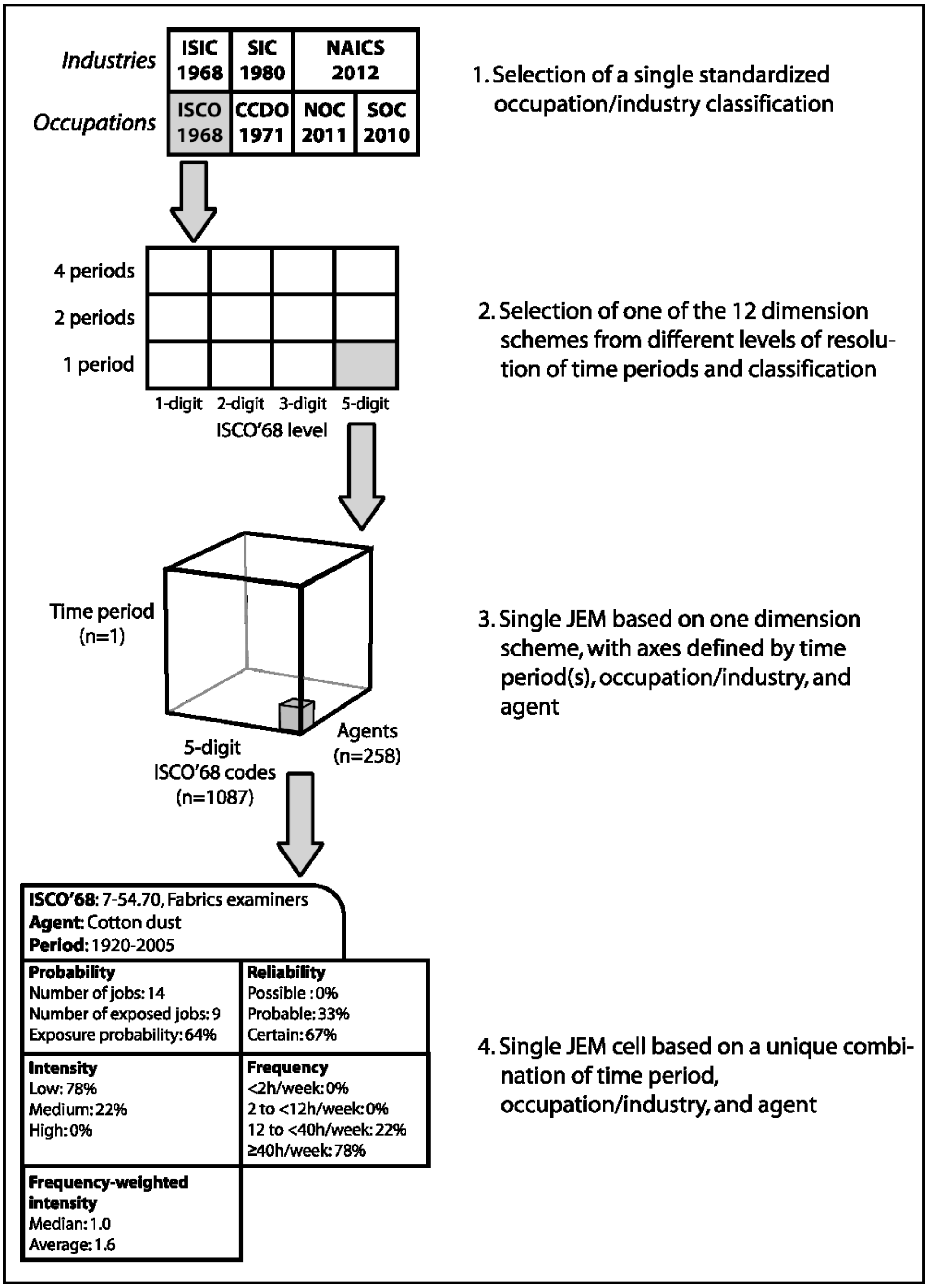
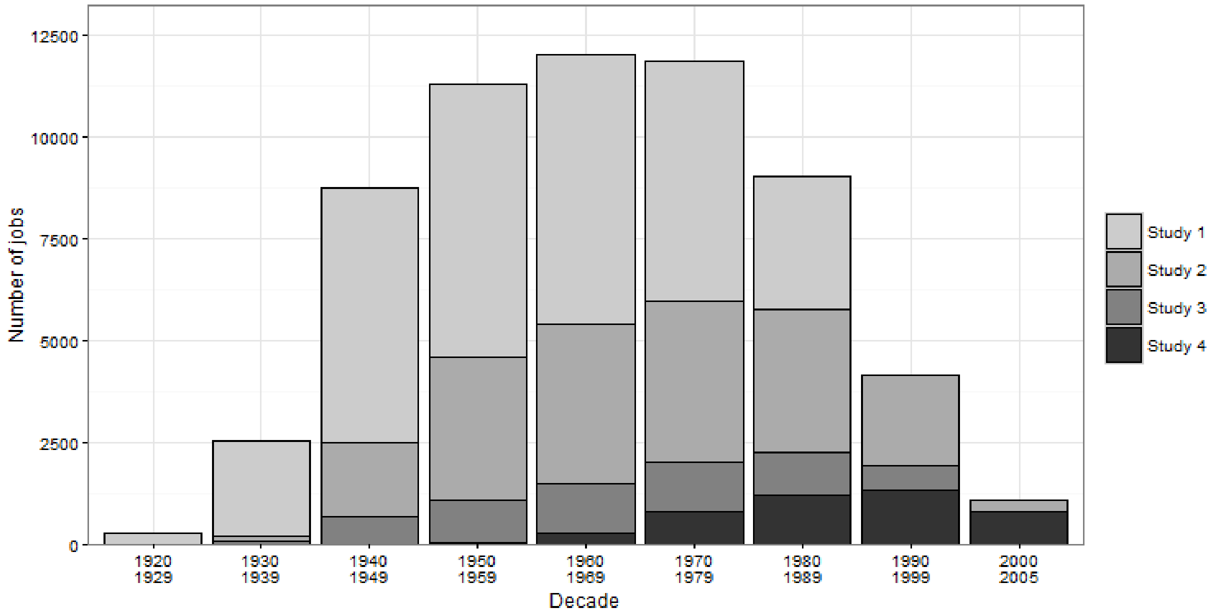


Figure 2. Number of jobs in the pooled exposure database by decade of employment[1], stratified by study



Since a job with a period of employment covering more than one decade was included in each time period category, the cumulative total is greater than 31,673

Chapitre 4. Development of the CANJEM job exposure matrix: Bayesian modelling of occupational exposures assigned by experts to over 30,000 jobs spanning 1930-2005

Development of the CANJEM job exposure matrix: Bayesian modelling of occupational exposures assigned by experts to over 30,000 jobs spanning 1930-2005

Jean-François Sauvé [1,2], Jack Siemiatycki [2,3], Marie-Élise Parent [2,3,4], Marie-Pierre Sylvestre [2,3], Jérôme Lavoué [1,2]

1. Université de Montréal, School of Public Health, Department of environmental and occupational health, Montréal, Québec, Canada
2. Centre de recherche du CHUM, Montréal, Québec, Canada
3. Université de Montréal, School of Public Health, Department of Social and Preventive Medicine, Montréal, Québec, Canada
4. INRS-Institut Armand-Frappier, Université du Québec, Laval, Québec, Canada

Jean-François Sauvé a contribué de façon majeure à la préparation des données, à leur analyse et à l’interprétation des résultats, ainsi qu’à la rédaction et à la préparation du manuscrit.

4.1 Abstract

Background: The CANJEM job-exposure matrix compiles expert evaluations of 31,673 jobs from four population-based case-control studies conducted in Montreal. For each job, experts had derived indices of intensity, frequency and probability of exposure to 258 agents. CANJEM summarizes the exposures assigned to jobs into cells defined by occupation/industry, agent, and period. Some cells may, however, be less populated than others, resulting in uncertain estimates. This paper describes a modelling framework to refine the estimates of sparse cells by drawing on information available in adjacent cells.

Methods: Bayesian hierarchical logistic and linear models were used to estimate the probability of exposure and the geometric mean (GM) of frequency-weighted intensity (FWI) in cells, respectively. The hierarchy followed the Canadian Classification and Dictionary of Occupations (CCDO) classification structure, allowing for exposure estimates to be provided across occupations (7-digit code), unit groups (4 digits) and minor groups (3 digits). The models were applied to lead compounds, formaldehyde, wood dust, silica and benzene, and four periods, adjusting for the study from which jobs were evaluated.

Results: The models allowed for the estimation of probability and FWI for all cells across the three levels of the CCDO classification, with estimates from sparsely populated cells pulled towards the average of the higher-level group. Overall, the effect of shrinkage towards the group mean was significant below 5 job/cell, moderate from 5 to 9 jobs/cell, and negligible at 10 jobs/cell or greater. For FWI, the model adequately pulled the estimates towards the group mean for all levels, whereas a systematic shift towards lower probability was found for the 3-digit group estimates, with median decrease of 0.8% compared to the descriptive estimates. The more

recent studies were associated with lower FWI for wood dust, and lower probability for benzene (1950-1969 only) and lead compounds; however, no overall trend in between-study differences emerged.

Conclusions: Albeit based on a small number of agents, the modelling framework for FWI appears to be a suitable approach to refine current CANJEM estimates. For probability, the models could be improved by methods adapted to the large number of cells with no exposure.

4.2 Introduction

Retrospective occupational exposure assessment in population-based studies requires the reconstruction of detailed work histories. Since obtaining relevant quantitative industrial hygiene measurement data is generally unfeasible, it has traditionally relied on indirect methods involving questionnaires and interviews, followed by expert review (Gérin *et al.*, 1985), or job-exposure matrices (JEMs) (Teschke *et al.*, 2002). A JEM is a cross-tabulation of occupations (or industries) and agents, with each unique combination of these two dimensions representing a cell with an exposure estimate. JEMs allow for an automatic attribution of exposures to jobs for which the occupation and/or industry title is available. In contrast, the expert method provides job-specific estimates of exposure that may vary between jobs sharing the same occupation. This, however, comes at a cost of greater complexity, time and manpower required for its implementation (Goldberg and Imbernon, 2002; Siemiatycki, 2007).

We recently reported on the construction of CANJEM, a general population JEM derived from a database of 30,000 jobs evaluated by experts in the context of four population-based case-control studies of cancer conducted in Montreal, Canada since the 1980s (Sauvé *et al.*, 2017). CANJEM was created by summarizing the exposures assigned to individual jobs (where each job represents an occupation held by a subject for at least 6 months) for 258 chemical and physical agents into strata of occupation or industry (available in several standardized classifications) and employment period. Both the occupation/industry and the period dimensions are defined over several levels of resolution, from specific occupations/industries to broader categories, or from a single global period to a stratification into 4 shorter periods. Each cell provides a descriptive summary of the exposures assigned to jobs according to their probability, reliability, intensity and frequency, and frequency-weighted intensity (FWI).

Since CANJEM was based on a finite sample of jobs held by subjects in the four studies conducted over a 25-year period, the quantity of information varies across cells, an issue especially acute for less prevalent occupations or industries. Increasing the resolution of the occupation/industry groups and periods also implies that the finite set of jobs is distributed across a larger number of categories, thereby further decreasing sample size per cell. One way to obtain a more precise estimate of exposure for a cell based on few jobs (*e.g.*, truck mechanics) is to simply use the estimate for the group at a lower level of resolution (*e.g.*, motor vehicle mechanics) pooling jobs across the nested occupations. This may represent a useful approach when the exposure profile in one occupation (or industry) is comparable to the other occupations within the same group. On the other hand, this could introduce bias if the exposure profile of the broader group is not indicative of the exposure in a specific occupation despite the increase in precision. Hierarchical models represent an alternative approach that could provide a compromise between the unbiased, but less precise information of cells at finer resolution, and the more precise, but also potentially biased information of coarser resolutions. The use of these models structured by the occupation/industry systems allows for cells based on a few data points to draw information from other, more populated cells associated with similar occupations within a broader group.

Hierarchical models have been applied in occupational exposure assessment to account for similarities in exposure profiles among workers, job titles or facilities (Friesen *et al.*, 2006; Toti *et al.*, 2006; Portengen *et al.*, 2016). Other examples include combining a generic JEM with measurement data to estimate quantitative exposure levels by occupation, where the occupations were grouped by their categorical JEM rating (Peters *et al.*, 2011; Friesen *et al.*, 2012; Peters *et al.*, 2016). Hierarchical models have also been used by our group in the evaluation of lung cancer risk for 184 agents by pooling information across agents sharing similar chemical characteristics and/or prior evidence of carcinogenicity (Momoli *et al.*, 2010).

In this paper, we created a Bayesian modelling framework to refine the estimates of sparse CANJEM cells by building on the information contained in cells of similar occupations, and explored the impact of this framework on final JEM estimates.

4.3 Methods

The Montreal case-control studies

Study populations

CANJEM summarizes data from four population-based case-control studies in Montréal, Canada. Study 1 (conducted 1979-1986) investigated 19 different sites of cancer among men aged 35-90 years (3,726 cancer patients and 533 population controls) (Siemiatycki *et al.*, 1987). Study 2 (1996-2001) was a study of lung cancer and included males and females aged 35-75 years (1203 cases and 1513 population controls) (Pintos *et al.*, 2012). Study 3 (1996-1997) was a study of breast cancer and included women aged 50-75 years (608 cases and 667 cancer controls) (Labrèche *et al.*, 2010). Study 4 (2000-2004) was a study of glioma and meningioma tumours and represented the Quebec and Ontario portions of the multi-centric INTEROCC study (Lacourt *et al.*, 2013), and included men and women aged between 30 and 59 years of age (218 cases and 414 population controls).

Occupational exposure assessment

The expert approach to exposure assessment described in Gérin *et al.* (1985), was developed during Study 1 and applied in subsequent studies. Briefly, complete occupational histories including job titles, employment duration, tasks performed, work environment and conditions and product and equipment use were collected from questionnaires and extensive face-to-face interviews with subjects, or proxy respondents. A team of trained chemists and industrial hygienists reviewed each job description, blind to the subject's case/control status, to assign standardized job and industry titles and to assess exposures to a predefined list containing approximately 300 chemical physical and biological agents. Exposure was rated by its intensity (low, medium, high), its frequency (hours per week) and the experts' level of confidence, or reliability, in the assessment (possible, probable, definite). A list of occupational circumstances corresponding to each intensity level was also devised for several agents which served as benchmarks in standardizing the exposure assessment. Each job was evaluated by two experts, and consensus was reached to resolve divergences in ratings

Exposure information in CANJEM

The occupational histories and exposure data associated with 31,673 jobs served as the foundation of CANJEM, with 15,067 jobs from Study 1 (47.6%), 10,371 from Study 2 (32.7%), 3,510 from Study 3 (11.1%) and 2,725 from Study 4 (8.6%). A detailed description of CANJEM can be found in Sauvé *et al.* (2017), and the exposure information available can be consulted freely at www.canjem.ca.

CANJEM dimensions

The information of jobs was summarized in CANJEM into three dimensions: agents, occupations/industries and periods. The agents axis includes 258 agents. The occupation/industry dimension is available in 7 standard classification schemes. For any of the 258 agents estimates of exposure can be obtained at several resolution levels of the selected classification, from broader groupings (*e.g.*, service occupations) to the most detailed categories (*e.g.*, waiters). The period dimension is available in three levels of resolution: a single global period (1930-2005), 2 periods (1930-1969 and 1970-2005), and 4 periods (1930-1949, 1950-1969, 1970-1984 and 1985-2005). Jobs with an employment period spanning two or more adjacent periods within one resolution could contribute information on exposure for all corresponding periods. Exposure estimates can be extracted across different resolutions of the occupations/industries and periods. Increasing the resolution of cells (i.e., more periods, finer occupational codes) may result in higher specificity in exposure levels, but also results in an estimate based on a smaller sample size.

Exposure indices of CANJEM cells

The exposure profile of jobs in each cell is represented by five indices: probability, reliability, intensity, frequency and frequency-weighted intensity (FWI) of exposure. Probability represents the proportion of jobs exposed among all jobs in the cell. Reliability is the relative proportion of jobs with possible, probable and definite ratings. Similarly, intensity of exposure presents the relative proportion of exposed jobs across the low, medium and high ratings, and frequency the relative proportion of jobs exposed <2 hours, 2 to <12 hours, 12 to <40 hours, and 40 hours per week or more. Lastly, FWI is a continuous index representing the intensity of exposure averaged over a 40-hour workweek, computed by multiplying the intensity ratings with the frequency of exposure relative to a baseline of 40 hours. In computing FWI, weights of 1, 5 and 25 were assigned to the low, medium and high intensity categories, respectively (Lavoué *et al.*, 2012). An FWI of 1 may thus be interpreted as exposure at low intensity for 40 hours per week, at medium intensity for 8 hours per week, or at high intensity approximately 1.5 hours per week. FWI in CANJEM cells is represented as the median value across exposed jobs.

Model development

General framework

A hierarchical modelling approach based on the structure of the occupational/industrial classification was used, in which exposure estimates could be provided by one model for all cells across the resolutions of a classification system. Cells from each period were modelled separately, allowing for the estimation of distinct time trends in exposure. Therefore, for a given combination of agent, period and exposure index, a single model provides estimates for cells across all resolutions of the selected occupational or industrial classification.

Furthermore, although all studies relied on the same general data collection and exposure assessment framework, shifts in the definition of exposure indices, refinement of questionnaires and accrual of experience may have caused differences in exposure estimates for a comparable situation. This would have been the most important for the time gap between the first study conducted by the group in the early 1980s, and the other studies which were conducted some 15 years later. To account for potential shifts in exposure coding, the models included a binary variable separating the older Multisite study from the others.

Application of models

The models were developed for two exposure indices, probability and FWI, and applied to CANJEM defined by four periods and the 1971 Canadian Classification and Dictionary of Occupations (CCDO) (Department of Employment and Immigration, 1971). This classification is structured with four hierarchical levels: 2-digit major groups, 3-digit minor groups, 4-digit unit groups and 7-digit occupations, the latter featuring 7907 unique codes. The major group strata being deemed too broad, the models only included the 3, 4 and 7-digit levels. The models were applied to five agents (lead compounds, formaldehyde, wood dust, crystalline silica and benzene) to encompass a diversity of physical forms. All available cells in CANJEM were used in the modelling, without restriction on sample size, since cells based on a single job could still provide information on exposure for higher-level groups.

The model for FWI was applied to the individual exposed jobs separately for each agent and period. Linear models were applied to the log-transformed FWI values to estimate the geometric mean (GM) FWI by occupational group based on the structure shown below.

$$\ln(FWI_{hijk}) \sim \beta_{Study=site-specific} \times X_{study} + \beta 3d_h + b4d_{hi} + b7d_{hij} \qquad \mathbf{1}$$

Where $\ln(FWI_{hijk})$ is the log-transformed FWI value of the k^{th} job in the j^{th} 7-digit group in the i^{th} 4-digit group in the h^{th} 3-digit group.

The 3-digit groups were entered as fixed effects in the model. The 4-digit groups ($b4d_{hi}$ in equation 1) were entered as a random-effects nested within 3-digit groups ($\beta 3d_h$), and the 7-digit groups ($b7d_{hij}$) were nested within the 4-digit groups. Lastly, the term for study was entered as a binary variable in the models applied to the two middle periods only (1950-1969 and 1970-1984) where the job histories overlapped the most. For the other periods, 73% of the data for period 1930-1949 came from Multisite whereas 95% of the data for period 1985-2005 came from the more recent studies.

For probability of exposure, which is expressed as a proportion, a logistic model was used along with the same core structure used to model FWI. The number of exposed jobs in a cell ($Nexp$) was modelled as a binomial distribution defined by the proportion of jobs exposed (π) and the total number of jobs evaluated ($Ntot$) (equation 2). The logit of π was assumed to follow a normal distribution based on the mean of the 7-digit group and the level of the study variable (when applicable).

$$Nexp_{hijk} \sim Binomial(\pi_{hijk}, Ntot_{hijk}) \qquad \mathbf{2}$$

Where $Nexp_{hijk}$ represents the number of exposed jobs in the cell for the jth 7-digit occupation and the kth study level. The logit of π was then modeled according to equation 3, allowing for the concurrent estimation of probability across the 3, 4 and 7-digit groups.

$$logit(\pi_{hijk}) = \beta_{Study=site-specific} \times X_{study=k} + \beta 3d_h + b4d_{hi} + b7d_{hij} \qquad \mathbf{3}$$

Where π_{hijk} represents the estimated proportion of exposed job of the kth study level in the jth 7-digit group in the ith 4-digit group in the hth 3-digit group.

The models were fitted in a Bayesian framework, with weakly informative prior distributions placed over the model parameters, using JAGS 3.4.0 software (Plummer, 2003). The appendix presents additional details on the prior distributions and computational methods used.

Development of predictions for CANJEM cells

The hierarchical model structure allowed for predictions to be made for the probability of exposure or the geometric mean (GM) of FWI for all cells across the three levels of the CCDO classification for one combination of agent and period. One important feature of hierarchical models is the borrowing of information across the data by shrinking the more imprecise estimates in the direction of the higher level group estimate. Exposure estimates for cells with few observations will tend to be pulled more heavily towards the mean of the higher-level group (*e.g.*, towards the 4-digit unit group for occupations), to a greater extent when their (unshrunk) estimates differ markedly from the group mean. On the other hand, cells with more observations would be less affected. This shrinkage allows an increase in precision in estimates while applying some level of bias towards the estimate of the higher level group (Greenland, 2000). Another potential benefit of this approach is that the estimates of cells of higher-level groups may better reflect the distribution of average exposure across the nested cells, compared to an average level weighted by the sample size of cells using a descriptive approach.

As an illustration consider a group of motor vehicle mechanics as one level, with two subgroups forming the lower level: automobile mechanics, with a FWI of diesel exhaust of 0.5 based on 50 jobs (equivalent to 20 hours exposed at low intensity), and heavy truck mechanics, with one exposed job and a FWI of 25 (40 hours at high intensity). The resulting overall estimate for mechanics would therefore be mainly based on jobs from automobile mechanics. Provided this distribution of mechanics jobs reflects the distribution in the base population (*i.e.*, higher proportion of automobile mechanics), the estimate would accurately represent the overall exposure profile for "mechanics". However, the estimate for truck mechanics would be pulled towards the overall estimate for mechanics, *i.e.,* closer to 0.5 than 25, itself driven by automobile mechanics. If an FWI of 25 actually reflects the true exposure of truck drivers in the population, this pulling effect is undesirable. On the other hand, if the exposure of truck drivers in the population is actually lower and the very high exposure for the single observation available in our database is due to random sampling, the pooling of the data would provide a better estimate.

Since it is not possible to discriminate between these two scenarios for every situation that may arise in CANJEM, we conducted an evaluation of the impact of sample size on the robustness of the estimates to shrinkage. This was done with the aim of finding a compromise value allowing for some, but not extreme, shrinkage for using the results of a cell.

The inclusion of the study in the models meant that predictions could be made for a cell for a situation reflecting only the Multisite study, only the site-specific, or a combination of both. The latter scenario was retained, with weights of 75% for site-specific studies and 25% for Multisite applied in the predictions for CANJEM cells. We used a higher weight for the more recent site-specific studies since experts had more experience with the coding approach and access to a larger pool of information to reconstruct past exposures. This adjustment was also performed for periods where the study term was omitted from the models; details for the weighting of the study on the predictions are presented in the Appendix.

4.4 Results

Descriptive statistics of the exposure data

The total number of jobs available, the corresponding number of cells by CCDO level, and the number of exposed jobs by agent for each of the four periods are presented in Table I. The total is greater than the total number of jobs (31,673) since jobs could be included in more than one period. Lead compounds had the highest proportion of exposed jobs for all periods except for 1985-2005, where formaldehyde ranked highest. The number of 7-digit cells with at least one exposed job varied from 94 (benzene, 1985-2005) to 616 (lead, 1950-1969).

Modelling

Between-study differences in exposure

The associations between site-specific studies and the probability and GM of FWI in cells relative to those of the Multisite study are presented in Table II. For probability, the largest difference was found for lead compounds: jobs from site-specific studies were far less likely to be exposed compared to those from the Multisite study, with odds ratios (OR) of 0.28 for 1950-1969 and 0.23 for 1970-1984. Compared to a hypothetical cell with a probability of 10% in Multisite, the corresponding probability in site-specific studies would be 3.1% for 1950-1969 and 2.5% for 1970-1984, respectively. The differences were comparatively smaller among the other combinations of agents and periods where the ORs generally leaned closer to 1. The influence of site-specific studies on the GM of FWI of cells was expressed as a relative percentage relative to a reference of 100% for Multisite (Lavoué *et al.*, 2006). Site-specific studies were associated with FWI levels on average 75-80% of those in Multisite for lead and silica, and 50% for wood dust. In the case of formaldehyde and benzene, the FWI levels of jobs were comparable between the two studies.

Predicted probability and GM of FWI of cells

To illustrate the distribution of the information on exposure in cells across the levels of the classification, Figure 1 presents the predicted probability and GMs of FWI (along with 90% credible intervals (CI)) for exposure to formaldehyde (where the between-study differences were negligible) among cells nested in the minor group of Fabricating, assembling and repairing occupations, wood products (CCDO 854) for the period 1970-1984. Approximately half of all jobs or exposed jobs were associated with the occupation of cabinetmakers (CCDO 8541-110), while most of the other occupations were based on 1 or 2 jobs. The predicted probability of cells that had 0% or 100% of jobs exposed were pulled away from these two extremes. For FWI, the estimate for Laminating-press tenders was particularly sensitive to the influence of other groups, from a value of 5 based on a single job, shrunk to a predicted GM of 0.91.

Effect of sample size on the sensitivity to shrinkage

Among the occupations listed in Figure 1, the estimated FWI for Laminating-press tenders was particularly impacted by the model due to its small sample size and large value relative to the other occupations. Another illustration of the influence of cell sample size on shrinkage of the estimates is presented in Figure 2, adapted from Raper (2013), taking as an example exposure to formaldehyde among all 7-digit cells within the major group 855/856 (Fabricating, assembling and repairing occupations: Textile, fur and leather products) for the period 1950-1969. The 62 nested 7-digit cells were categorized in three groups by cell sample size (exposed jobs for FWI): fewer than 5 jobs, 5 to 9 jobs and 10 jobs or more. In each panel, the descriptive

GMs of 7-digit groups are plotted on a diagonal line. The dots represent the GMs of the 4-digit cells corresponding to each 7-digit group with shapes denoting the different 4-digit cells. The ends of the arrows represent the predicted estimates of the 7-digit groups. Longer arrows correspond to a larger difference between the descriptive and predicted estimates, and thus a stronger shrinkage effect. For both probability and FWI, the level of shrinkage tends to decrease from the leftmost panel (<5 jobs) to the rightmost one (≥10 jobs), with a marked decrease in sensitivity with a sample size of at least 5 jobs per group.

Predicted vs descriptive probability and FWI

Figure 3 presents a comparison of the distribution of the descriptive and predicted probability (Figure 3a) and FWI (Figure 3b) across all cells stratified by level of the CCDO classification, using exposure to formaldehyde in period 1970-1984 as an illustration. For FWI, the models pulled the very low or very high estimates towards the overall average for cells at the occupation (7-digit) and unit group (4-digit) levels, where no systematic differences were observed in one direction or another. No appreciable difference was found at the minor group level (3-digit) due to their inclusion as fixed effects in the model. A different pattern emerged for probability: while the predicted estimates of 7-digit cells contained fewer extreme values (i.e., 0% or 100%), the shrinkage for the 3-digit cells went systematically in the direction of lower probability of exposure, with a median decrease of 1.7% in the predicted probability of cells relative to their descriptive estimates (interquartile interval 0-6.4%). This pattern was consistent throughout the analyses, where the median decrease in the predicted probability of 3-digit cells ranged 0-4.3% (median 0.8%) across the agents and periods.

4.5 Discussion

In this paper, we developed a Bayesian modelling framework to improve estimates of probability and FWI in CANJEM cells based on a small number of jobs, by integrating information on exposure available in other related occupations based on the structure of a standardized classification. This results in estimates that are a compromise between the level of information on exposure specific to jobs evaluated in a cell, and the information available in other cells within the same broader occupational group. In a recent application of CANJEM to evaluate exposure to pesticides (Zeng *et al.*, 2017), cells based on fewer than 10 jobs were deemed less informative, and the estimates of cells in broader occupation groups were used instead. While this conservative approach may increase the precision of the estimates, it can also result in less specific information on exposure being used. The modelling framework presented here aimed to provide estimates that represent a compromise between the less precise, but more specific information of finer occupations, and the more precise, but less specific information of broader groups.

Predicted probability and FWI

Effect of sample size and classification structure on shrinkage

The borrowing of information on exposure between cells in the models was organized by the hierarchical structure of the occupational classification, which implies some level of exchangeability in exposure across occupations. Taking the example of Figure 1, this would suppose that the exposure to formaldehyde is *a priori* comparable between the various occupations of cabinet and wood furniture makers in the absence of exposure data specific to an occupation. The estimates of individual cells would then draw on the information available within the larger pool of cabinet and wood furniture makers in the models, which would increase the precision by adding some amount of bias compared to a purely descriptive estimate. The trade-off is greater for cells with an outlying estimate relative to the other cells, and with a smaller sample size (and thus with a higher uncertainty). This shrinkage property can be useful in facilitating the inclusion of groups with a small sample size in an analysis that could otherwise result in unstable estimate. However, the possibility a large bias outweighing the increased precision of an estimate was a concern when applying this framework to the context of CANJEM due to the challenge of differentiating the lack of compatibility in the exposures between occupations from random variation.

While the grouping of occupations in the CCDO classification is primarily based on similarities in the work performed and, to some extent, in the services provided or goods produced (Employment and Immigration Canada, 1989), it may not always reflect similarities in exposure profiles for all occupations and for all agents. For example, Service station attendants (CCDO 5145) belong to the minor group of Sales occupations, commodities (CCDO 513/514), which also includes Supermarket clerks and Pharmaceutical representatives. The exposure profile to benzene or engine emissions for service station attendants may lean closer to other occupations associated with motor vehicles such as mechanics and driving occupations. These might represent more comparable groups to draw information from, as opposed to other sales occupations. However, defining alternative groupings of occupations based on exposure profiles rather than similar tasks and activities would require a significant amount of expert judgment, and would have to be conducted for each of the four classification systems for occupations and the three classifications for industries used in CANJEM. Moreover, the similarities in the exposure profiles of occupations or industries to a given agent may not translate to other agents. For example, grouping service station attendants with automobile mechanics might not be reasonable with regards to exposure to asbestos or degreasing agents. The dependencies between categories would thus need to be defined by experts for each of the 258 agents, which would not be feasible with limited resources available The evaluation of shrinkage showed that overall, cells with fewer than 5 jobs for probability, or exposed jobs for FWI, tended to be quite sensitive to the shrinkage effect (Figure 2), while cells with 5 to 9 jobs were more robust and shrinkage was negligible for those with at least 10 data points. A sample size of 5 jobs, while an arbitrary threshold, may represent a reasonable starting point in using the estimate of a cell that accounts to some extent for the information available in similar occupations, without being overly sensitive to shrinkage towards the group mean. While the structure of the classification may not always be representative of the distribution in exposures, the impact of potential misspecification is therefore limited when using a sample size of at least 5 jobs per cell, and avoids defining alternative schemes to group occupations based on exposures. Regarding FWI, this evaluation was based on the number of exposed jobs in the cell, and not the total number of jobs. Cells with a large number of jobs may thus have an estimate for FWI based on only one or two exposed jobs. Applying the FWI estimates should therefore also account for the probability of exposure together with considerations of sample size.

Overall trends in the predicted probability and FWI of cells

The distribution of the predicted GM of FWI of cells went in the expected direction, where the more outlying estimates (high or low) were pulled towards the mean, suggesting the suitability of the models in pooling the exposure information across CANJEM cells. On the other hand, the model for probability resulted in systematically lower predicted values at the highest hierarchical level (3-digit minor group). The exploration of alternative models, which are presented in more detail in the Appendix, showed that pattern arises from a lack of compatibility of logistic models with a data structure predominantly composed of (7-digit) cells with no exposure, which ranged from 74% to 93% across the agents and periods. The use of a linear model resulted in a distribution of predicted probability analogous to FWI in Figure 3. While its usefulness is limited since the predictions may fall outside the range of 0-100%, it nevertheless showed that the hierarchical structure can be suited for probability. While the amplitude of the systematic shift towards lower probabilities was limited, the use of models adapted for zero counts, such as zero-inflated binomial regression in a hierarchical framework (Hall, 2000), might constitute an improved strategy to model the probability of exposure of CANJEM cells. The adaptation of these models to the multiple hierarchical levels of the classifications and the unbalanced structure of the data would require however further methodological development. Further improvements could also extend the models to the ordinal indices of intensity and frequency of exposure. However, as with probability of exposure, the unbalanced structure of the data, such as the high proportion of jobs exposed at low intensity relative to the medium and high ratings, also represents a challenge to the application of traditional modelling approaches.

Between-study differences in exposure

This analysis also represented an opportunity to examine and adjust for potential differences between studies in the exposures assigned to jobs by the experts stratified by cell. Overall, no clear trend was apparent across agents and the between-study differences were moderate with three notable exceptions, where site-specific studies were associated with lower FWI for wood dust and lower probability of exposure for benzene (1950-1969 only) and lead compounds. For probability, it is possible that some of the differences observed represent another effect of the application of the models to a large number of cells with no exposure. Other potential sources of differences between source studies are the increased experience and familiarity of the team with the exposure assessment method over time and changes in the meaning benchmarks for the intensity categories and for the background environmental exposure level (Pintos *et al.*, 2012). Finally, the models provided a global estimate of the between-study differences across all cells. Further exploration to identify whether the elevated differences might be driven by specific occupations or exposure circumstances should help to shed some light on our observations regarding probability for lead and benzene, and wood dust for FWI.

The inclusion of a variable for the source studies in the model allowed us to weigh the relative influence of each study across all cells in the predictions. In contrast, the influence of the study on exposure estimates obtained using a descriptive approach might vary from cell to cell depending on the distribution of the jobs between the two study levels within each cell. We placed a higher weight to the more recent site-specific studies. The impact of this on the predictions of cells would be influenced by the magnitude of the between-study difference, from relatively large for probability of exposure to lead compounds, to negligible for formaldehyde. The application of the models to a broader list of agents would provide more information as to the amplitude of the differences between studies, and whether some patterns emerge across agents that would be indicative of more generic changes in exposure assessment.

4.6 Conclusion

In this paper, we presented a modelling framework to refine the estimates of less populated cells in CANJEM using the hierarchical structure of an occupational classification system, providing systematic method to share information on exposure between cells of similar occupations. The models applied to the index of FWI appeared to have adequately weighted the influence of cells between the hierarchical levels on the final estimates. Their application to probability was however suboptimal, likely due to the large number of cells with no exposure, and further developments could help refine the model predictions. Some differences in exposure were found between the source studies of jobs comprised in CANJEM, although no systematic trend across agents was observed. The framework presented here could by used for the analysis of other existing databases of past expert evaluations.

4.7 Tables and figures

Table I. Total number of jobs per time period, and corresponding number of exposed jobs by agent.

	Time period			
	1920-1949	**1950-1969**	**1970-1984**	**1985-2005**
Overall				
Number of jobs available[1]	9444	17147	13450	6405
Number of 7-digit occupations[2]	1743	2408	2082	1289
Number of 4-digit unit groups[2]	436	469	461	392
Number of 3-digit minor groups[2]	79	81	81	80
Number of exposed jobs by agent (%)				
Lead compounds	1568 (16.6%)[3]	2837 (16.5%)	1898 (14.1%)	336 (5.2%)
Formaldehyde	965 (10.2%)	1960 (11.4%)	1522 (11.3%)	663 (10.4%)
Wood dust	1007 (10.7%)	1525 (8.9%)	916 (6.8%)	341 (5.3%)
Silica	807 (8.5%)	1480 (8.6%)	907 (6.7%)	277 (4.3%)
Benzene	571 (6.0%)	1055 (6.2%)	541 (4.0%)	145 (2.3%)

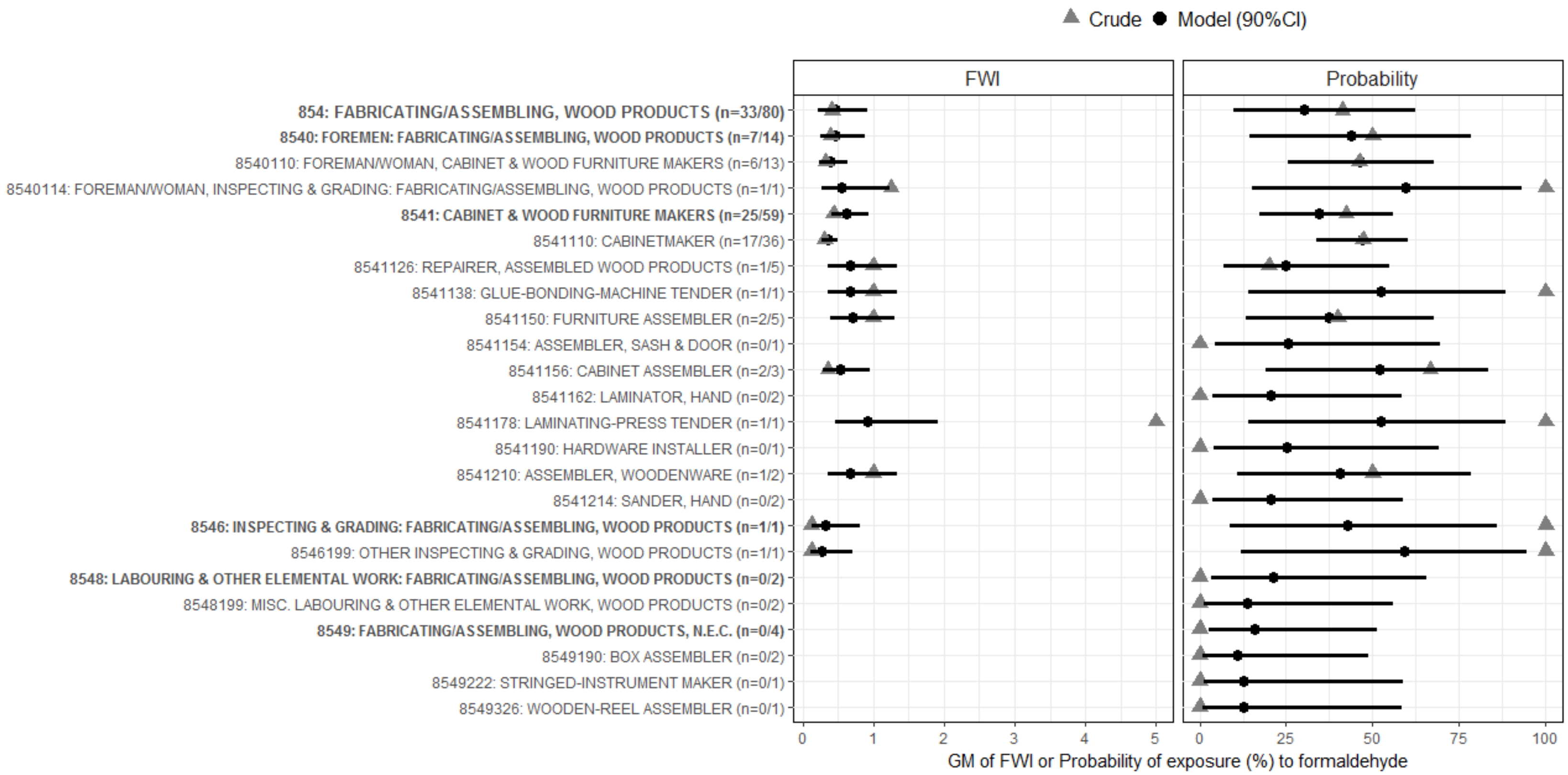
1. Since jobs could be present in two or more adjacent periods, the total is greater than the number of jobs used in the construction of CANJEM (n=31,673)
2. Number of groups with at least one job available
3. Percentage of exposed jobs relative to the total number of jobs assessed within the time period

Table II. Relative influence of site-specific studies on the probability and FWI of exposure of cells relative to Multisite, stratified by time period and agent.

		Probability		FWI	
	Effect measure	**Odds ratio (90% CI)** (reference: Multisite = 1)		**Relative index of exposure (%) (90% CI** (reference: Multisite = 100%)	
	Period[1]	**1950-1969**	**1970-1984**	**1950-1969**	**1970-1984**
Agent	**Lead compounds**	0.28 (0.26-0.32)[2]	0.23 (0.20-0.26)	81 (74-88)[3]	76 (70-84)
	Formaldehyde	0.93 (0.83-1.05)	1.05 (0.92-1.21)	118 (109-128)	108 (98-118)
	Wood dust	1.27 (1.10-1.47)	1.19 (1.00-1.42)	51 (46-58)	45 (38-52)
	Silica	0.72 (0.62-0.84)	0.93 (0.78-1.11)	72 (64-82)	80 (68-94)
	Benzene	1.93 (1.66-2.24)	1.17 (0.96-1.43)	102 (88-117)	105 (87-126)

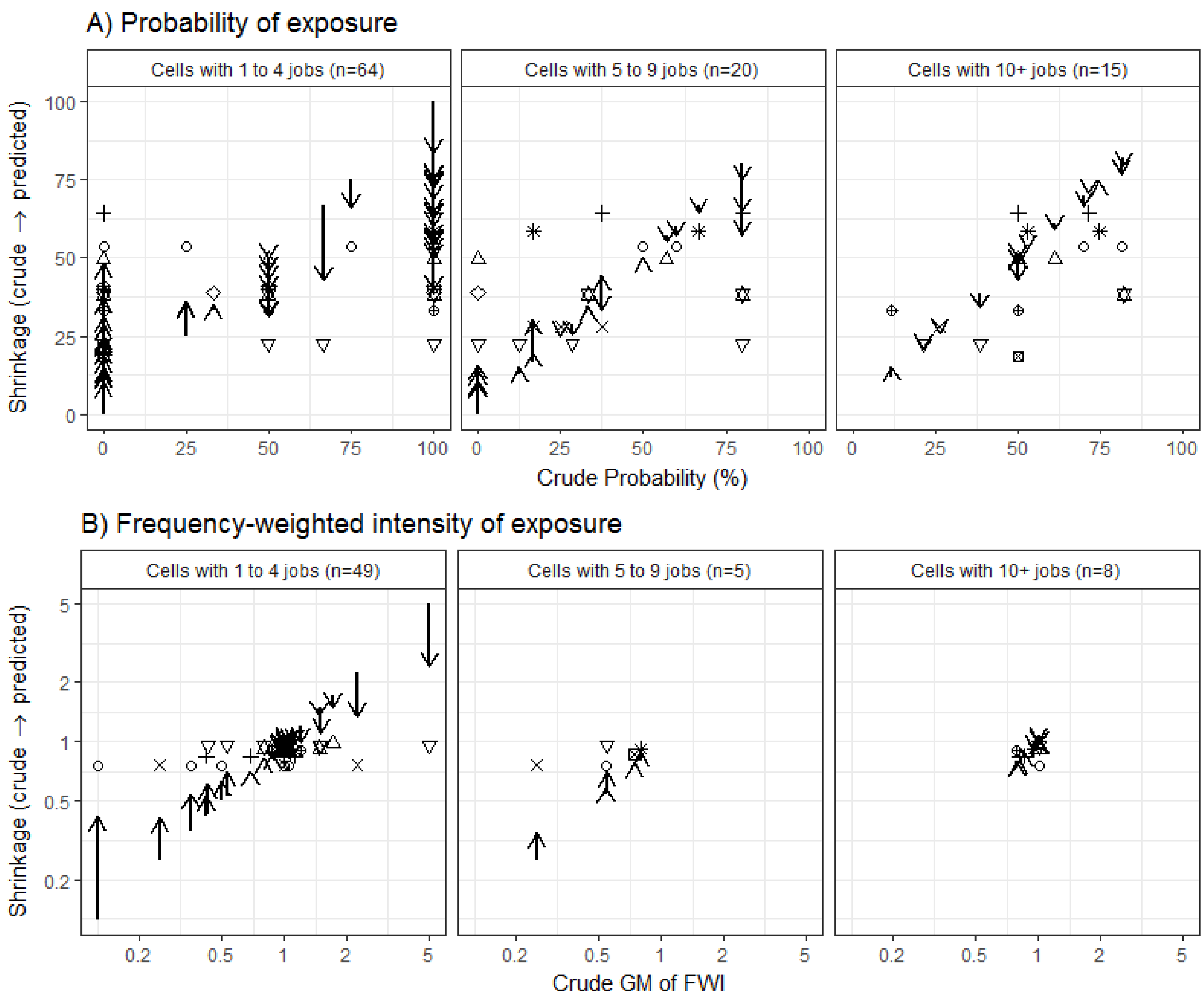
1. The inclusion of study in the models was limited to periods 1950-1969 and 1970-1984.
2. Odds ratio and 90% credible interval for site-specific studies, relative to a reference of 1 for the Multisite study.
3. Relative index of exposure (Lavoué et al., 2006) and 90% credible interval for the influence of site-specific studies on the GM of FWI in cells, relative to the reference of 100% for the Multisite study.

Figure 1. Comparison of the descriptive and predicted probability and GM for FWI for exposure to formaldehyde among cells nested in CCDO minor group 854 (Fabricating, assembling and repairing occupations, wood products), period 1970-1984



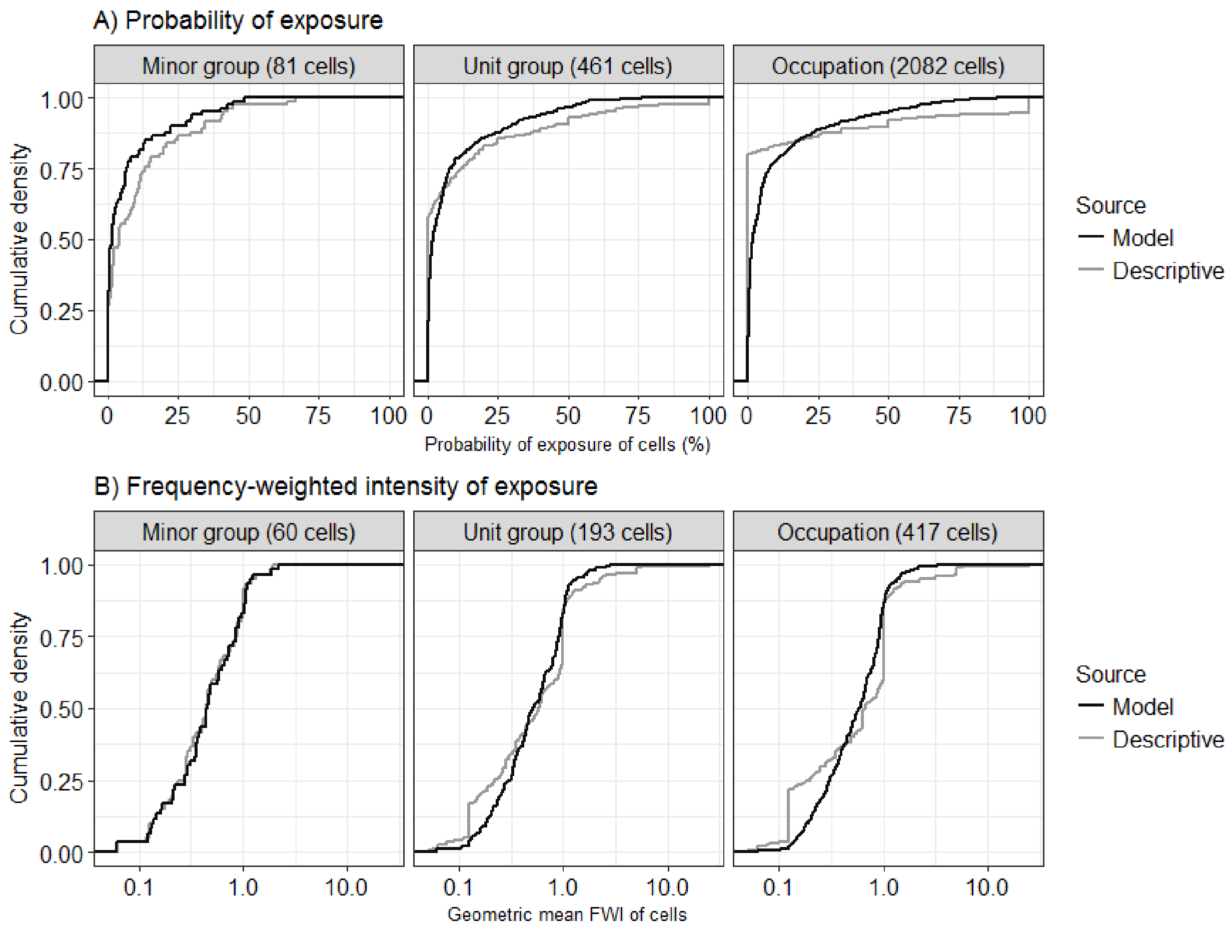
Numbers in parentheses denote the number of exposed jobs and the total number of jobs in a cell, respectively. CI: Credible interval

Figure 2. Illustration of the effect of cell sample size on the shrinkage of the estimates for probability and FWI of exposure to formaldehyde, period 1950-1969, among 7-digit cells (n=99 for probability, n=62 for FWI) within major group 855/856 (Fabricating, assembling and repairing occupations: Textile, fur and leather products).



Each individual arrow represents one 7-digit cell, while the shapes denote the 4-digit cell associated with each 7-digit cell. The start of the arrow represents the crude estimate obtained using a descriptive approach, and the end of the arrows represents the predicted estimate shrunk in the direction of its 4-digit cell.

Figure 3. Comparison of empirical cumulative distribution functions of the descriptive and predicted estimates of cells across all three levels of the CCDO classification, for the probability and FWI of exposure to formaldehyde, period 1970-1984.



4.8 References

Brooks SP, Gelman A. (1998) General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics; 7: 434-55.

Department of Employment and Immigration. (1971) Canadian classification dictionnary of occupations, vol. 1. Ottawa, ON: Department of Employment and Immigration.

Employment and Immigration Canada. (1989) Canadian Classification and Dictionary of Occupations – Guide, 9th edition. Ottawa: Employment and Immigration Canada, Occupational and Career Information Branch.

Friesen MC, Coble JB, Lu W *et al.* (2012) Combining a Job-Exposure Matrix with Exposure Measurements to Assess Occupational Exposure to Benzene in a Population Cohort in Shanghai, China. Ann Occup Hyg; 56: 80-91.

Friesen MC, Macnab YC, Marion SA *et al.* (2006) Mixed Models and Empirical Bayes Estimation for Retrospective Exposure Assessment of Dust Exposures in Canadian Sawmills. Ann Occup Hyg; 50: 281-88.

Gelman A, Shirley K. (2011) Inference from Simulations and Monitoring Convergence. In Brooks, S, Gelman, A, Jones, GL, Meng, X-L. Handbook of Markov chain Monte Carlo, Boca Raton, Fla.: CRC Press,.

Gérin M, Siemiatycki J, Kemper H *et al.* (1985) Obtaining occupational exposure histories in epidemiologic case-control studies. J Occup Med; 27: 420-6.

Goldberg M, Imbernon E. (2002) The use of job exposure matrices for cancer epidemiology research and surveillance. Archives of Public Health; 60: 173-85.

Greenland S. (2000) Principles of multilevel modelling. Int J Epidemiol; 29: 158-67.

Hall DB. (2000) Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study. Biometrics; 56: 1030-39.

He H, Wang W, Hu J *et al.* (2015) Distribution-free inference of zero-inflated binomial data for longitudinal studies. Journal of Applied Statistics; 42: 2203-19.

Labrèche F, Goldberg MS, Valois M-F *et al.* (2010) Postmenopausal breast cancer and occupational exposures. Occup Environ Med; 67: 263-69.

Lacourt A, Cardis E, Pintos J *et al.* (2013) INTEROCC case--control study: lack of association between glioma tumors and occupational exposure to selected combustion products, dusts and other chemical agents. BMC Public Health; 13: 340.

Lavoué J, Pintos J, Van Tongeren M *et al.* (2012) Comparison of exposure estimates in the Finnish job-exposure matrix FINJEM with a JEM derived from expert assessments performed in Montreal. Occup Environ Med; 69: 465-71.

Lavoué J, Vincent R, Gérin M. (2006) Statistical Modelling of Formaldehyde Occupational Exposure Levels in French Industries, 1986-2003. Ann Occup Hyg; 50: 305-21.

Momoli F, Abrahamowicz M, Parent ME *et al.* (2010) Analysis of multiple exposures: an empirical comparison of results from conventional and semi-bayes modeling strategies. Epidemiology; 21: 144-51.

Peters S, Vermeulen R, Portengen L *et al.* (2011) Modelling of occupational respirable crystalline silica exposure for quantitative exposure assessment in community-based case-control studies. J Environ Monit; 13: 3262-68.

Peters S, Vermeulen R, Portengen L *et al.* (2016) SYN-JEM: A Quantitative Job-Exposure Matrix for Five Lung Carcinogens. Ann Occup Hyg; 60: 795-811.

Pintos J, Parent M-É, Richardson L *et al.* (2012) Occupational exposure to diesel engine emissions and risk of lung cancer: evidence from two case–control studies in Montreal, Canada. Occup Environ Med; 69: 787-92.

Plummer M. (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd international workshop on distributed statistical computing.

Portengen L, Linet MS, Li G-L *et al.* (2016) Retrospective benzene exposure assessment for a multi-center case-cohort study of benzene-exposed workers in China. J Expos Sci Environ Epidemiol; 26: 334-40.

Raper S. (2013) Visualizing shrinkage. http://www.coppelia.io/2013/08/visualising-shrinkage/ (August 31, 2013).

Sauvé J-F, Labrèche F, Richardson L *et al.* (2017) CANJEM: a general population job exposure matrix based on past expert assessments of exposure to over 250 agents. Manuscript in preparation

Siemiatycki J. (2007) Investigating cancer risks related to asbestos and other occupational carcinogens. Occup Environ Med; 64: 500-1.

Siemiatycki J, Wacholder S, Richardson L *et al.* (1987) Discovering carcinogens in the occupational environment. Methods of data collection and analysis of a large case-referent monitoring system. Scand J Work Environ Health; 13: 486-92.

Teschke K, Olshan AF, Daniels JL *et al.* (2002) Occupational exposure assessment in case-control studies: opportunities for improvement. Occup Environ Med; 59: 575-94.

Toti S, Biggeri A, Baldasseroni A. (2006) A hierarchical Bayesian model for a variability analysis of measurements of occupational n-hexane exposure in Italy. Statistical Modelling; 6: 175-85.

Zeng F, Lerro C, Lavoué J *et al.* (2017) Occupational exposure to pesticides and other biocides and risk of thyroid cancer. Occup Environ Med [In press] doi: 10.1136/oemed-2016-103931

4.9 Appendix

Implementation of the Bayesian models

Weakly informative priors were used over all model parameters. Normal priors with mean 0 and variance 1000 were used for the coefficients of 3-digit groups (β_{3d}) and study, when present in the model. Prior distributions for the between-occupation and between-unit group variances, as well as for the within-occupation variance (FWI only), were uniform distributions on the scale of the standard deviation bounded between 0.001 and 100, with the lower bound set to avoid a standard deviation of zero (corresponding to an infinite precision). Each model for FWI was fitted using 12 MCMC chains with 75,000 iterations each, discarding the first 25,000 iterations, for a total of 600,000 iterations used for inference. Since the models for probability applied to all CANJEM cells (and not only those with exposure jobs), 275,000 iterations per chain were used, with the first 25,000 discarded and keeping the results of one out of each 10 iteration (total of 300,000 iterations kept). Convergence in the simulations was assessed using the Brooks-Gelman-Rubin statistic, or Rhat (Brooks and Gelman, 1998), where a value close to 1 indicates convergence for a given parameter.

Computations were made on the computer cluster Briarée from the Université de Montréal, managed by Calcul Québec and Compute Canada, using JAGS version 3.4.0 (Plummer, 2003). The number of MCMC chains matched the number of cores on a node of the Briarée cluster, allowing for the computations to be run simultaneously. Fitting the model for probability of exposure for all four periods took approximately 4 hours for each agent, while models for FWI took between 25 to 50 minutes depending on the agent. For the models for FWI, convergence was reached for all parameters with Rhat values lower than 1.1 (Gelman and Shirley, 2011) for all combinations of agents and periods. For probability, there remained fewer than 5% of the model parameters with Rhat values above 1.1.

Adjustment for the source studies of jobs in the predictions for periods 1930-1949 and 1985-2005

In adjusting the estimates for the study of jobs in the predictions, weights of 25% for Multisite and 75% for site-specific were used. Equation 4 provides an example for the adjustment of the predicted GM of FWI in a cell for the periods 1950-1969 and 1970-1984 in which the variable for the source study was included in the models:

$$GM_{pred} = exp\left(\mu + 0.75 \times \beta_{site-specific}\right) \qquad \mathbf{4}$$

Where GM_{pred} is the predicted GM of FWI for a cell, weighted by study; μ is the coefficient for the cell; 0.75 is the weight given to site-specific study; and $\beta_{site-specific}$ is the coefficient for site-specific studies for the period.

For consistency, and to avoid the potential distortions highlighted previously, the predictions for the two periods without the study variable were also adjusted using the same weights to the two levels, and the coefficient for site-specific studies from the adjacent period (*e.g.,* for 1930-1949, the median value of the posterior distribution for the study coefficient from 1950-1969), accounting for the proportion of jobs from site-specific studies in each cell. Equation 5 provides an example of this adjustment on the predictions:

$$GM_{pred} = exp\left(\mu + [0.75 - \frac{Nexp_{site-specific}}{Nexp_{total}}] \times \beta_{site-specific}\right) \qquad \mathbf{5}$$

Where GM_{pred} is the predicted GM of FWI for a group weighted for the studies; μ is the coefficient for the cell; 0.75 is the weight given to site-specific studies; $Nexp_{site-specific}$ is the number of exposed jobs from site-specific studies in the cell; $Nexp_{tot}$ is the total number of jobs in the cell, and $\beta_{site-specific}$ is the coefficient for site-specific studies taken from the adjacent period.

Alternative models for probability of exposure

The comparisons between the distribution of the exposure estimates obtained from a descriptive approach to those predicted from the model coefficients exhibited a different pattern between probability and FWI when stratified by level of the CCDO classification. In particular, the predicted probability of 3-digit cells weas systematically lower relative to the crude proportion of exposed jobs. A likely explanation stems from a lack of compatibility of the logistic model to a dataset comprised of a large proportion of cells with no exposed job. To this end, we investigated the use of alternative Poisson and linear regression models for probability, along with logistic models.

A first alternative evaluated was to consider the number of exposed jobs in a cell as arising from a Poisson distribution (equation 6).

$$Nexp_{hijk} \sim Poisson(\lambda_{hijk}) \qquad \mathbf{6}$$

Since the aim was to model the probability of exposure in the cell rather than the expected number of jobs, the rate λ_{hijk} was standardized by cell sample size by including the logarithm for the total number of jobs in the cell as an offset in the model.

$$\log(\lambda_{hijk}) = \beta_{Study=site-specific} \times X_{study=k} + \beta 3d_h + b4d_{hi} + b7d_{hij} + \log(\, Ntot_{hijk}) \qquad \mathbf{7}$$

Where λ_{hijk} represents the estimated number of exposed jobs of the kth study level in the jth 7-digit group in the ith 4-digit group in the hth 3-digit group, and $Ntot_{hijk}$ the total number of jobs in the cell.

The second approach examined the use of a linear regression model. Compared to the logistic and Poisson regression models, which were applied to 7-digit cells, the application of the linear models was based on the exposed/unexposed status of individual jobs, with values of 0 or 1 taken for an unexposed or an exposed job, respectively (equation 8).

$$exposed_{hijk} \sim \beta_{Study=site-specific} \times X_{study} + \beta 3d_h + b4d_{hi} + b7d_{hij} \qquad \mathbf{8}$$

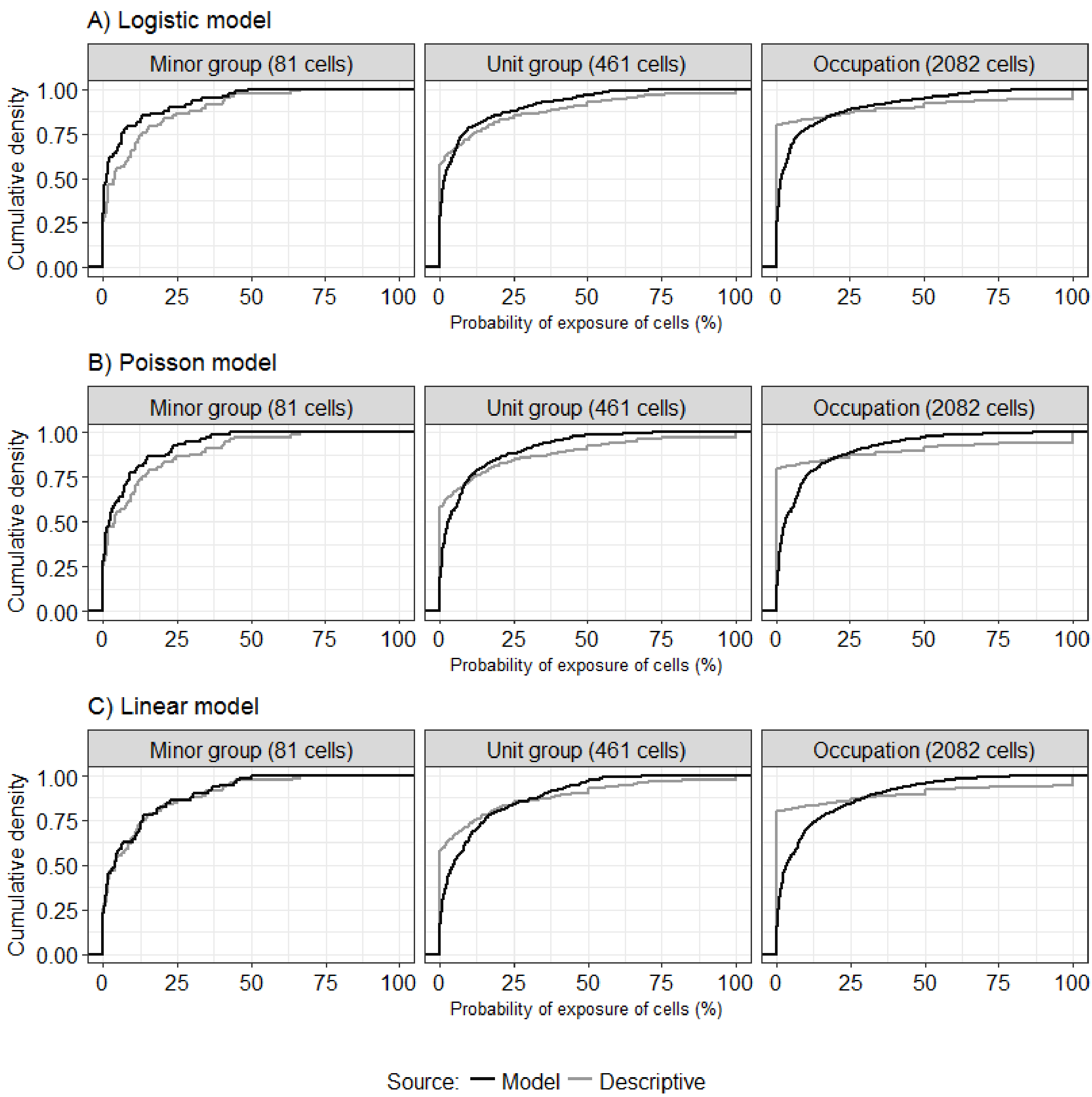
Where $exposed_{hijk}$is the binary exposed/unexposed status of the kth job in the jth 7-digit group in the ith 4-digit group in the hth 3 digit group. Due to the model being applied to individual jobs as opposed to 7-digit cells, this required an additional parameter for the within-occupation variance, akin to the model for FWI.

These models were fitted using the same set of prior distributions and MCMC iterations as the logistic regression model. In addition, these models were only applied to the probability of exposure to formaldehyde, for which the difference between studies observed was the smallest across the five agents evaluated, in order to minimize the influence of the study on the trends observed in the predicted probability of cells.

Supplementary figure 1 extends the comparison of the distribution of the descriptive and predicted probability of exposure of cells across the three levels of the CCDO classification shown in Figure 3 to the three types of models examined. The results obtained with the Poisson regression model did not differ markedly from those of the logistic model, where the systematic trend of lower predicted probability remained at the minor group level. On the other hand, this pattern in the predicted probability of cells was not found with the linear model, which resulted in a distribution of probability in 3-digit cells that was comparable to the one obtained using a descriptive approach, analogous to the trends found for the modelling of FWI.

While the linear model was less sensitive to the imbalance in the distribution of probability resulting from the large proportion of cells with no job exposed, its application is not suited to data in the form of proportions that are constrained to a narrow range of possible values (*i.e.*, between 0% and 100%). This exercise did however provide some clues as to the relative lack of performance observed with the logistic model. The application of zero-inflated Poisson (ZIP) or binomial (ZIB) models within a hierarchical framework (Hall, 2000) may constitute an alternative approach to account for the large number of cells with no exposed jobs while being better suited to the modelling data in the form of proportions. These essentially consist of a mixture of two models, one for "structural" zeros and the other for "sampling" zeros resulting from sampling variability (He *et al.*, 2015). The former may correspond for example to a situation where an agent is never present in an occupation or industry; the number of exposed jobs would therefore always be zero. For the latter, some jobs within an occupation may be exposed in the population, but none were exposed among the sample of jobs collected during the studies. ZIP and ZIB models can therefore disentangle these two processes generating a distribution with a large mass at zero, allowing for the application of a logistic model to the subset of groups that excludes those with structural zeros – in our case, limited to cells for which some exposure would possible.

Supplementary figure 1. Comparison of empirical cumulative distribution functions of the descriptive and predicted probability of exposure to formaldehyde in period 1970-1984 of cells across all three levels of the CCDO classification, obtained using logistic, Poisson and linear regression models



Chapitre 5. Development of quantitative estimates of wood dust exposure in a Canadian general population job-exposure matrix based on past expert assessments

Development of quantitative estimates of wood dust exposure in a Canadian general population job-exposure matrix based on past expert assessments

Jean-François Sauvé [1,2], Hugh W Davies [3], Marie-Élise Parent [2,4,5], Jérôme Lavoué [1,2]

1. Université de Montréal, School of Public Health, Department of environmental and occupational health, Montréal, Québec, Canada
2. Centre de recherche du CHUM, Montréal, Québec, Canada
3. University of British Columbia, School of Population and Public Health, Vancouver, British Columbia, Canada
4. INRS-Institut Armand-Frappier, Université du Québec, Laval, Québec, Canada
5. Université de Montréal, School of Public Health, Department of Social and Preventive Medicine, Montréal, Québec, Canada

Jean-François Sauvé a contribué de façon majeure à la préparation des données, à leur analyse et à l'interprétation des résultats, ainsi qu'à la rédaction et à la préparation du manuscrit.

5.1 Abstract

Objectives: The CANJEM general population job-exposure matrix summarizes expert evaluations of 31,673 jobs from four population-based case-control studies of cancer conducted in Montreal, Canada. Intensity in each CANJEM cell is represented as relative distributions of the ordinal (low, medium, high) ratings of jobs assigned by the experts. We aimed to apply quantitative concentrations to CANJEM cells using Canadian historical measurements, taking exposure to wood dust as an example.

Methods: Wood dust measurements came from the Canadian Workplace Exposure Database (CWED). We selected personal and area samples in occupations (2011 Canadian National Occupational Classification) with a non-zero exposure probability in CANJEM in period 1930–2005 (minimum 10 samples/occupation in CWED). Concentrations were modelled with sampling duration, year and type, source database and proportion of jobs at medium and high intensity in JEM cells (fixed effects), and occupations (random effects).

Results: 5170 samples from 31 occupations spanning 1981–2003 were retained. Estimated geometric mean (GM) concentrations for a cell with all jobs at medium or high intensity were respectively 1.3 and 2.3 times higher than a cell with all jobs at low intensity. An overall trend of -5%/year in exposure was observed. Predicted GMs for 8 hours, breathing zone and year 1989 for CANJEM cells associated with exposure ranged 0.49–1.67 mg/m3.

Conclusions: The model provided estimates of wood dust concentrations for any CANJEM cell with exposure, even for those without measurements by using the calibrated intensity ratings. This framework could be implemented for other agents represented in both CANJEM and CWED.

5.2 Introduction

Retrospective occupational exposure assessment in community-based studies has traditionally relied on expert judgment of individual job descriptions (Gérin *et al.*, 1985) or on job-exposure matrices (JEMs) that provide exposure estimates by occupation or industry titles (Hoar *et al.*, 1980). Exposures assigned to individual jobs or JEM cells are often expressed categorically, such as none, low or high exposure, mainly due to the limited availability of relevant historical workplace measurement data to estimate quantitative exposure levels. Some exceptions exist, such as the FINJEM general-population JEM from Finland (Kauppinen *et al.*, 1998) and Matgéné from France (Févotte *et al.*, 2011), that provide quantitative average exposure levels for several agents.

Measurement data collected by regulatory agencies for compliance monitoring have been identified since the 1990s as a potential useful source of exposure information for population studies (Stewart and Rice, 1990; Lippmann, 2011). These include the Integrated Management Information System (IMIS) in the United States (Lavoué *et al.*, 2013) and MEGA in Germany (Koppisch *et al.*, 2012). These both contain over 2 million measurements dating back to the 1970s and have been applied to some extent in support of exposure assessment efforts for epidemiological studies (eg., Teschke *et al.*, 1999; Kendzia *et al.*, 2013). Other large national databases include COLCHIC and SCOLA from France (Mater *et al.*, 2016), NEDB from the United Kingdom (Burns and Beaumont, 1989) and SIREP from Italy (Scarselli *et al.*, 2008). The application of these databases in exposure assessment for population studies has historically been limited due to concerns regarding the non-random selection of workers or industries monitored (Olsen *et al.*, 1991), the inconsistent quality of descriptive information (Hall *et al.*, 2002) and the lack of data for some occupations and industries. Recently, a framework initially developed for an industry-based study (Wild *et al.*, 2002) to systematically apply measurement data for retrospective exposure assessment by calibrating a semi-quantitative JEM, has been adapted to population studies (Peters *et al.*, 2011; Friesen *et al.*, 2012). The JEM represents a prior opinion of the presence or absence of exposure for an occupation, and of the average exposure level of jobs expressed as categorical rating. The combination of these two sources allows for the estimation of quantitative exposure levels by occupation and for each categorical JEM rating. It provides a mechanism for addressing data gaps by using the calibrated JEM ratings as quantitative estimates for occupations not represented in the measurement data. Moreover, the JEM ratings can also be used to adjust for potential biases in the measurement data for the remaining occupations (Peters *et al.*, 2011).

We recently reported on the creation of CANJEM, a general population JEM summarizing expert evaluations of over 31,000 jobs collected during four large case-control studies in Montreal, Canada (Sauvé *et al.*, 2017). CANJEM features three dimensions: occupation or industry, agent (n=258), and employment period. When the experts had considered a job exposed to an agent, they rated the intensity of exposure on a three-point scale (low, medium, high). For CANJEM, the information on exposure intensity is represented by the relative proportion of jobs exposed at low, medium and high level. One issue is that these categories do not directly correspond to quantitative concentration levels. This could be addressed using historical workplace measurement data, such as those from the recently constituted Canadian Workplace Exposure Database, or CWED (Hall *et al.*, 2014; Peters *et al.*, 2015; Hon *et al.*, 2017). CWED comprises exposure measurements for over 300 agents, mostly collected for compliance purposes by provincial governmental agencies in Canada since the 1960s.

The objective of this paper is to present an approach to obtain quantitative retrospective estimates of exposure for CANJEM cells by combining the distribution of semi-quantitative ratings within occupations in the JEM with the measurement data contained in CWED using statistical models. This paper focuses on exposure to wood dust to initiate the development of the quantitative dimension of CANJEM.

5.3 Methods

Data sources

CANJEM

CANJEM is a publicly-available (www.canjem.ca) general-population JEM constructed from a database of exposure evaluations of jobs performed by experts over the course of four population-based case-control studies conducted in the Montreal metropolitan region in Canada (Sauvé *et al.*, 2017). Overall, information on exposure to 258 agents from 31,673 jobs held by 8,760 subjects between 1930 and 2005 was used to build CANJEM. In each study, interviews and semi-structured questionnaires were used to collect lifetime occupational histories for each subject covering job title, company name, main tasks and activities, general work environment and use of protective equipment. A team of trained experts reviewed each job description to assign occupation and industry codes from standardized classifications, and to assign exposures to a predefined list of agents, including wood dust, using the method described in Gérin *et al.* (1985). The experts did not discriminate between dusts from softwood, hardwood and/or allergenic species, thus "wood dust" in these studies (and by extension CANJEM) covers all species. When an agent was considered to be present in the workplace at levels above those found in the general environment, the experts described the exposure profile with three indices: reliability, or level of confidence in the assessment (possible, probable, definite), intensity (low, medium, high) and frequency (hours per week exposed). Benchmark occupations or processes associated with each intensity rating were used to guide the experts. Benchmarks for wood dust included construction carpenters (low intensity), sawmills (medium) and belt sanding (high)

(Vallières *et al.*, 2015). These did not constitute fixed rules and could be modulated based on the job descriptions.

The pooled job histories and exposure data was summarized in CANJEM by three dimensions: occupation or industry, period and agent. The occupation/industry axis is available in four Canadian and international standardized classifications for occupations and three for industries, and in several levels of resolution within each classification. The time axis is defined by three levels of resolution, from one global period (1930-2005) split into two or four periods. The 258 agents form the third dimension. Each CANJEM cell represents a unique combination of occupation/industry, period and agent for which at least one job was evaluated and summarizes the exposures of jobs with five indices: probability (or proportion of jobs exposed), reliability, intensity, frequency, and frequency-weighted intensity (FWI). Intensity represents the relative proportion of jobs exposed at low, medium and high levels. FWI is a continuous index multiplying intensity with frequency of exposure (hours per week exposed) averaged over a 40-hour week. In computing the FWI of jobs, the low, medium and high intensity categories were assigned weights of 1, 5 and 25, respectively. Thus, exposure at low intensity for 40 hours or at medium intensity for 8 hours both represents an FWI of 1. While these relative weights may have varied from one agent to another, the 1-5-25 represented the best overall estimate across agents according to the experts (Lavoué *et al.*, 2012). The median FWI value across exposed jobs in the cell was used as the estimate in CANJEM.

Canadian Workplace Exposure Database (CWED)

CWED is an ongoing project initiated in 2008 to compile existing exposure data stored in electronic databases and in hard copy to create a national measurement database for Canada

(Hall *et al.*, 2014). Serving as a centralized repository of historical measurement data from Canadian workplaces, CWED has been used for the surveillance of exposure to carcinogens in the Canadian population in the CAREX Canada project (Peters *et al.*, 2015), in estimating the burden of occupational cancer in Canada (Demers *et al.*, 2014), and in documenting historical trends in exposure to isocyanates (Hon *et al.*, 2017).

CWED contains approximately 500,000 measurements from 350 agents, collected for different purposes such as compliance or routine monitoring and research. About 80% of the data originates from the provinces of British Columbia and Ontario, the latter from the Medical Surveillance (MESU) database (Lubek, 2011), with the remainder from other provinces and federal administrations. While the measurements in CWED were collected between the 1960s and the 2010s, most were taken from the mid-1970s to the late 1990s. Ancillary information includes sampling date, contaminant sampled, company name, sampling duration, sample type (area or personal), sampling method, occupation title (2006 National Occupation Classification for Statistics, or NOC-S) (Statistics Canada, 2007) and industry title (2002 North American Industry Classification System, or NAICS) (Statistics Canada, 2003)).

Data preparation

Selection of measurement data

A total of 6569 wood dust entries were retrieved from 6 individual databases of 4 provinces (BC, Manitoba, Ontario, Saskatchewan) and covering the period 1978-2012. The samples contained a mixture of softwood (n=1673, 24%) or hardwood (n=2062, 30%), allergenic (n=1389, 20%) and non-allergenic (n=1452, 21%) species, or were unspecified or mixtures. We excluded samples with missing occupation title, sampling duration or year, a sampling duration under 15 minutes, or a concentration not expressed in milligrams per cubic metre. Area samples, comprising approximately 30% of the data, were retained while those with unreported sample type were excluded. Reporting of sampling methods was partial. Most samples from BC used the WorkSafeBC method 1150 for "Particulate (Total) in Air" using a 37 mm cassette. Methods for the Ontario data were not reported but the occupational exposure limits (OELs) are also based on "total" dust using comparable methods. Some samples had a method reported for nitrous oxides or using gas chromatography and were excluded. Information on the limit of detection (LOD) was available for 2841 samples (41% of total) from BC. LOD values were 0.1 mg/sample before 1994, and 0.05 mg/sample from 1994 onwards. These values were assigned to the remaining samples based on the year. 248 samples with a missing, negative or zero concentration value, and 497 samples with a concentration lower than the imputed LOD were flagged as non-detects for a total of 745 samples (11% of total). We excluded samples from the BC Mines (n=2) and Manitoba Ministry of Labor Workplace Health and Safety Division (n=41) databases due to the small sample size. This resulted in an initial database of 5428 samples, originating from 3 sources: BC Laboratory Information Management System (LIMS) 1 (n=1723), BC LIMS 2 (n=167) and Ontario MESU (n=3538).

Linkage of CANJEM and CWED

We selected CANJEM defined by the 2011 NOC classification (Statistics Canada, 2012), the closest to the 2006 NOC-S classification in CWED. 4-digit 2011 NOC codes were assigned for each CWED measurement using official conversion tables (Statistics Canada, 2016b). When multiple codes were possible, we selected the occupation with a similar title (maximum difference of 5 characters); otherwise the occupation with the largest number of Canadians employed was selected. The following estimates from CANJEM cells in period 1930-2005 were then assigned to each measurement: number of jobs evaluated, number of jobs exposed, exposure probability, median FWI, and relative proportion of exposed jobs at low, medium and high intensity.

Statistical analyses

Statistical modelling

The model for combining the CWED measurement data with CANJEM was based on the framework described in Peters *et al.* (2011) and Friesen *et al.* (2012), in which log-transformed concentrations represent the outcome, and the predictors comprise the categorical ratings of cells as fixed effects and the exposure groups as random effects. Quantitative exposure levels can then be estimated for each categorical JEM rating as well as specific levels by individual exposure group. By using a random effect model, the estimates for groups with few measurements are pulled towards the overall mean, resulting in a comparatively higher influence of the JEM ratings on the predicted exposure levels relative to groups with more data points available (Wild *et al.*, 2002; Peters *et al.*, 2011). We adapted this structure to the multiple intensity categories in CANJEM cells using one term for the proportion of jobs at medium intensity and another for the proportion at high intensity. Other covariates were sample year, duration and type (area or personal) and source database, in the model structure shown below:

$$\ln(Y) = \beta_0 + \beta_{med}P_{med} + \beta_{high}P_{high} + \beta_{year}Year + \beta_{dur}Duration + \beta_{type}Type + \beta_{db}Database + Random_{NOC4d} + \varepsilon$$

Where

- $\ln(Y)$ was the natural log-transformed wood dust concentration

- β_0 : model intercept
- $\beta_{med}P_{med}$and $\beta_{high}P_{high}$: continuous variables for the relative proportions of jobs in CANJEM cells exposed at medium and high intensity levels, respectively
- $\beta_{year}Year$: continuous variable for sample year (standardized with 1978=0)
- $\beta_{dur}Duration$: continuous variable for log-transformed sampling duration (minutes)
- $\beta_{type}Type$: categorical variable for sample type (area or personal)
- $\beta_{db}Database$: categorical variable for the source database
- $Random_{NOC4d}$: random-effect term for occupation (4-digit NOC code), assumed to be distributed normally with mean 0 and variance σ_w^2.

The model coefficients were transformed into relative indices of exposure (RIE) (Lavoué *et al.*, 2006) to illustrate the influence of the variables as a percentage of increase or decrease on exposure relative to a reference level (taken as 100%).The model was fitted with a Bayesian approach using the JAGS 4.0.0 software (Plummer, 2003), using 12 Markov chain Monte Carlo (MCMC) chains of 13,750 iterations each, discarding the first 1250 iterations per chain and keeping one out of 2 of the remaining samples for inference (75,000 iterations total). The medians of the posterior distribution of the parameters were used as point estimates, with the 5th and 95th percentiles defining the 90% credible intervals (CI). Non-detected concentrations were treated as missing observations in the model, where they were imputed at each MCMC iteration from a distribution truncated at the limit of detection.

Prior distributions of fixed effect terms other than the intensity ratings were normal distributions of mean 0 and variance 5.7. In defining these prior distributions, we did not expect extremely large differences in exposure over time or between source databases, for instance. These priors represent a 50% interval of the RIEs covering the range 20%-500% (ie., from 5 times lower to 5 times higher than the reference). A larger prior variance of 46.6 was used for the intensity ratings to reflect the possibility of larger effects sizes on the exposure levels. The priors represented a 50% interval over the RIE covering the range 0.0001-10000% (ie., 100 times lower to 100 times higher than the reference of all jobs at low intensity). These priors were constrained so that the medium category was associated with higher exposure relative to low, and that the high category was associated with higher exposure relative to medium. This was done by first truncating the distributions to the positive domain, and setting β_{C_high} to a value greater than β_{C_medium}. Uniform(0,10) priors were used for the between and within occupation variances (log scale). Supplementary File 1 presents an example of the JAGS code used to fit the model.

Predictions

For each occupation with measurements available, we predicted the geometric mean (GM) of wood dust concentration using the following scenario: the year 1989 (median in dataset), duration of 480 minutes, personal sample type, relative proportion of measurements in each source database (shown in Table I), relative distribution of intensity ratings in the corresponding CANJEM cell, and the random intercept of the occupation ($Random_{NOC4d}$). Predicted GMs corresponding to a hypothetical cell with all jobs at low, medium or high intensity, were also made using the same combination of sample year, duration and type, and source database. We also predicted levels for all CANJEM cells with some exposure (regardless of measurements being available) based on their relative distribution of intensity ratings.

Sensitivity analyses

We conducted sensitivity analyses to evaluate the impact of alternative model specifications, exposure indices and data selection on the association between the intensity ratings and the wood dust concentrations. We first removed the constraint that forced a trend of higher exposure levels with higher intensity ratings in cells. Second, we evaluated a model with a single rating per cell by selecting the category with the most jobs. When two or three ratings had the same number of jobs, one category was randomly selected in each MCMC chain prior to fitting the model. Third, we tested a model with the log-transformed median FWI of cells as an alternative to the categorical intensity. Fourth, we applied different criteria on sample size, with a minimum of 1 or 5 samples per occupation, or restricted to occupations with probability ≥5%. Lastly, we performed analyses restricted to periods 1970-2005 or 1985-2005, using the CANJEM intensity estimates of cells specific to each period.

5.4 Results

Data selection

CANJEM contained 476 4-digit NOC cells for period 1930-2005. 198 cells (listed in Supplementary table 3) had at least one job exposed to wood dust, with a median probability of 6.3% (range 0.14-100%) The 5428 CWED samples initially retained corresponded to 78 occupations. Eight occupations (90 samples) had no exposed job in CANJEM and mainly concerned automobile and boat manufacturing, education and health care. 31 of the remaining 70 occupations with some exposure in CANJEM had at least 10 samples and were retained in the analysis (n=5170 samples). Woodworking machine operators (NOC 9437, n=1616) and Sawmill machine operators (NOC 9431, n=1127) had the largest number of measurements.

Descriptive statistics of the exposure data

Table I presents the number of samples, number of non-detects, GM and geometric standard deviation (GSD) of wood dust concentrations for selected categorical variables. The overall GM was 1.62 mg/m^3 (GSD 4.4), and 412 samples (8.0%) were non-detects. Sampling duration ranged from 15 to 690 minutes (median 195 minutes). Samples were collected between 1981 and 2003 (median 1989). Only 13% of samples were collected before 1985.

The number of samples, GM and GSD of wood dust concentrations by occupation, and the associated probability, intensity and FWI of exposure in CANJEM cells for period 1930-2005 are presented in Supplementary table I. NOC code 7321-Automotive service technicians, truck and bus mechanics and mechanical repairers (n=14) had the highest GM at 4.62 mg/m^3, followed by 8421-Chain saw and skidder operators (GM=4.59 mg/m^3, n=14) and 1241-Administrative assistants (GM 3.53 mg/m^3, n=13). Figure 1 presents the distribution of the exposure probability across the 198 CANJEM with at least one exposed job, differentiating between the 31 cells with at least 10 samples and the remaining 167 cells. Cells with a small exposure probability were less represented by measurements, while 13 of the 16 cells with a probability >60% had at least 10 samples. Figure 1 also displays the distribution of measurements by the probability of exposure of their respective CANJEM cells. Sixty-five percent of all measurements came from occupations with a probability greater than 90% in CANJEM. Among the 31 occupations, 6 cells (19%) had most jobs exposed at high intensity; 64% of all 5170 samples were associated with these 6 occupations.

Modelling

Estimated effects of variables on wood dust exposure levels

The point estimates and 90% credible intervals of the model coefficients are presented in Supplementary table II; Table II presents the RIEs for selected determinants included in the model. An annual decrease in exposure of 5%/year was found, while a 50% increase in sampling duration (eg., from 60 to 90 minutes) corresponded to a 26% reduction in exposure. Samples from the Ontario MESU and BC LIMS 2 databases were both associated with higher exposure compared to BC LIMS 1. For the association between the wood dust concentrations and the intensity ratings, using a cell with all jobs at low intensity as a reference, exposure levels were 1.3 times higher for a cell with all jobs at medium intensity, and 2.3 times higher for a cell with all jobs at high intensity.

Predicted exposure levels by occupation and intensity ratings

The predicted GM for 1989 corresponding to a cell with all jobs at low intensity was 0.75 mg/m^3 (90%CI 0.56-0.94 mg/m^3), compared to 1.02 mg/m^3 for all jobs at medium intensity (90%CI

0.81-1.31 mg/m^3) and 1.68 mg/m^3 for high intensity (90%CI 1.21-2.49 mg/m^3). In the absence of measurement data for an occupation, its estimate would therefore range between 0.75-1.68 mg/m^3, depending on its distribution of intensity ratings. The predicted GMs based on the relative distribution of ratings in cells of the 198 occupations are presented in Supplementary table III. The median GM across occupations was 0.87 mg/m^3 (interquartile interval 0.75-1.02 mg/m^3). For the 31 occupations with measurements, the predicted GMs based on the combined cell ratings and the occupation-specific mean from the model ranged from 0.49 mg/m^3 (2252-Industrial designers) to 1.67 mg/m^3 (7272-Cabinetmakers), representing a 3.4-fold difference compared to 2.3 for the calibrated ratings alone. The 10 occupations with the highest predicted GMs are listed in Table III, and mainly concerned woodworking and furniture manufacturing, logging and forestry, and carpentry. A Kendall correlation of 0.39 was observed between the GMs predicted from the ratings only and those combining ratings and occupation means.

Sensitivity analyses

Table IV presents the RIEs of categorical intensity ratings estimated with alternative model specifications and CANJEM cell information. The differences observed in the contrasts between categories were overall subtle. Removing the constraint on the order of the coefficients resulted in lower contrasts in the point estimates between categories and a greater uncertainty for the medium category. Lower contrasts also resulted from a less restrictive minimum sample size by occupation (≥1 or ≥5 samples). On the other hand, larger contrasts between intensity categories were found with a higher minimum exposure probability (5%) or using CANJEM estimates from period 1970-2005. Lastly, a positive trend was also observed with FWI, where a 50% increase in median FWI (e.g., from 1 to 5 days per week exposed at the same intensity level) translated to a 3.6% increase in GM concentrations (90%CI 0.1-7.3%) and a 1.7-fold difference between the highest and lowest predicted GMs of cells.

5.5 Discussion

In this paper, we modelled historical Canadian measurements to estimate quantitative word dust concentration levels and applied them to any cell in CANJEM for which at least one job was deemed exposed by the experts. The exposure levels assigned with the model represented a weighted estimate between the ratings of the cell and the occupation-specific random effect, with a greater weight for the former when the sample size was small. For the remaining occupations, quantitative concentrations were estimated indirectly by applying the calibrated ratings from the model to the relative distribution of intensity in cells.

Comparison of occupations represented in CWED and in CANJEM

There was a good concordance overall between the availability of measurements in an occupation and the presence of some exposure in CANJEM cells. Only 8 of the 78 occupations with measurements had an exposure probability of 0% in CANJEM, and these had a relatively small sample size with a median of 5 measurements per occupation. Measurements were also found for some occupations with a low probability of exposure. In some cases, such as Material handlers, Machinists and machining and tooling inspectors and Industrial painters, the presence of samples could be explained by the industrial sector such as sawmills, wood products manufacturing and boatbuilding. In other cases (eg., Administrative assistants and secondary school teachers), it may reflect unusual exposure circumstances, but this interpretation is speculative due to the lack of detailed contextual information (eg., sampling reason) accompanying the measurement data. However, most of the data was concentrated among occupations with a very high probability of exposure in CANJEM (eg., greater than 60%). In particular, Woodworking machine operators and Sawmill machine operators, which together represented 53% of all samples, had an exposure probability ≥95%. These occupations are also related to some of the industrial sectors (Sawmills, furniture and other wood product manufacturing) with the largest number of workers exposed to wood dust in Canada, after construction (CAREX Canada, 2015), which may explain the large number of measurements. There were comparatively fewer measurements in construction occupations, a trend also observed in CWED for other prevalent agents in this sector such as silica or diesel exhaust (Peters *et al.*, 2015). Lastly, while most occupations with samples had some exposure in CANJEM, the majority of occupations with a non-null probability of exposure in CANJEM were not represented in CWED. The use of a JEM can therefore provide a better indicator of the presence or absence of exposure across the spectrum of occupations in the population, compared to an exposure database alone.

Exposure levels by occupation and categorical intensity ratings

Compared to other studies combining measurement data with a generic JEMs based on a single rating per cell, the use of a distribution of intensity ratings in CANJEM cells allows for a greater variability in the calibrated estimates between cells when no measurement data is available. Across the 198 cells with at least one exposed job on CANJEM, predicted levels for 1989 varied all the way from 0.75 mg/m^3 (all jobs at low intensity) to 1.68 mg/m^3 (all jobs at high intensity) based on the calibrated ratings alone. The predicted GMs by occupation represent a product of the expert judgment of jobs in CANJEM cells, the exposure levels in the measurement data, and the relative sample size from the use of random effects. This resulted in a ratio of 3.4 between the smallest (0.49 mg/m^3) and largest (1.67 mg/m^3) predicted GMs by occupation for 1989, greater than the calibrated ratings alone (ratio of 2.3). In addition, integrating the CANJEM ratings led to the adjustment of some more outlying estimates compared to the measurements alone. For example, administrative assistants, with all exposed jobs at low intensity in CANJEM, had the 3rd highest descriptive GM based on 13 samples (3.53 mg/m^3), while its predicted GM from the model only ranked 16th (0.96 mg/m^3) among the 31 occupations, which suggests that the few measurements available may have represented unusual exposure circumstances for this occupation.

The ratios between GMs for all jobs at low, medium and high were on the order of 1, 1.3 and 2.3, which are smaller than the ratios of 1, 5 and 25 applied in the computation of FWI. The 1-5-25 scale represented an average estimate of these ratios across all 258 agents in CANJEM, and the results observed in this study suggests that this scale might not be adapted for wood dust specifically. The concentration of samples in occupations associated with higher intensity ratings may have also limited our ability to estimate larger differences in exposure between the low, medium and high categories.

The contrasts observed between the categories were however comparable to those observed in other studies combining expert ratings with measurement data (Peters *et al.*, 2011; Friesen *et al.*, 2012; Koh *et al.*, 2014; Peters *et al.*, 2016). For example, the predicted benzene concentrations by rating for 1980 in Friesen *et al.* (2012) were 2.5 mg/m^3 (low), 4.0 mg/m^3 (medium) and 7.0 mg/m^3 (high), corresponding to ratios of 1:1.6:3.0 between categories. In sensitivity analyses, the use of a single rating per cell in the model, based on most frequently assigned category, yielded lower contrasts between categories. This selection method did not differentiate between a cell with 51% of jobs at medium intensity and another with 100% at the same level. The use of relative distributions of ratings in the model, providing information on the within-occupation heterogeneity in intensity of jobs, may have therefore led to larger contrasts in exposure. Greater contrasts were also found in analyses restricted to occupations with a probability of exposure of at least 5%. Occupations with a very small probability also tended to have low exposure intensity; this change in the relative distribution may have contributed to the higher contrasts estimated. A similar pattern, albeit more subdued, was also found using CANJEM cells estimates for the period 1970-2005, but not for 1985-2005.

Determinants of exposure

The modelling of the wood dust measurement data also allowed for the assessment of factors potentially associated with exposure levels such as temporal trends. The annual decrease in exposure of 5%/year between 1978-2008 is comparable to trends observed in the United States (7%/year between 1979-1997) and the United Kingdom (8%/year between 1985-2005) (Teschke *et al.*, 1999; Galea *et al.*, 2009). The association between lower exposure levels with longer sampling durations has also been observed previously (eg., Lavoué *et al.*, 2006; Peters *et al.*, 2011; Kendzia *et al.*, 2017) and may result from the inclusion of periods without exposure during the monitoring period.

Regarding sample type, personal samples were associated with higher exposure levels relative to area samples by a factor of 50%. As discussed in Friesen *et al.* (2012), the relationship between personal and area measurements depends on factors such as the location of the sampling equipment relative to the source of exposure. While personal samples are preferred for this reason, we retained information from area samples, representing approximately 30% of the data, and adjusted for sample type in the model instead. Regarding differences between source databases, samples from Ontario's MESU were on average 1.4 times higher relative to BC LIMS

1. A similar difference was also found for BC LIMS 2 compared to BC LIMS 1, although based on only 160 samples. An opposite trend of higher levels in BC relative to Ontario was found in an analysis of isocyanate exposure levels in CWED (Hon *et al.*, 2017) that might be explained in part by differences in industries and regulations. For the former, measurements associated with sawmills tended to be more represented in the BC data, whereas MESU contained a higher proportion of data associated with manufacturing. However, most occupations included in this analysis were associated with a single sector (*eg.,* sawmill machine operators), which should account for these differences. As for differences in regulations, the current OELs for hardwood are 1 mg/m^3 for both provinces, but the OEL for softwood in BC of 2.5 mg/m^3 (non-allergenic species) is twice lower than in Ontario (5 mg/m^3), which may factor in the higher levels found for the latter. Lastly, the prior distributions for the effects of these variables on the exposure levels tended to be rather conservative, as the 90% credible intervals of the posterior distributions of the RIEs in Table II were all well within the range of 20-200%. The overall impact of these priors was therefore negligible.

Limitations

The individual databases forming CWED were not designed for secondary usage such as retrospective exposure assessment. Missing data was prevalent for several key variables such as the use of control methods or protective equipment that represented some of the factors used by the Montreal experts when assigning exposures based on job descriptions. The issue of missing data has also been reported previously by Hon *et al.* (2017) for isocyanates and by Hall *et al.* (2002) for softwood dust in the BC data. This also concerned sampling methods, such as the size fraction of dust particles collected. However, the OELs in BC and Ontario were based on total dust during the period covered by the data, which limits the possibility of the time trend being confounded by changes in the dust fraction sampled (*eg.,* from total to inhalable dust). The LOD was also missing for most of the data, and we used the values reported from samples from BC to populate the rest of the data. In doing so, we also used a conservative approach where reported concentration values below the LOD were considered non-detects. This may have led to an overestimation of the exposure levels since censoring was applied at a higher value than the result reported (with no information on the true detection status of the concentration), compared to an approach where these samples were considered detected.

Another undocumented factor that may influence exposure levels is the reason for sampling. A recent analysis of the IMIS database found that samples collected during follow-up inspections were associated with higher exposure, which may reflect a preferential selection of workplaces where exposure had previously been found (Sarazin *et al.*, 2016). The use of the CANJEM intensity ratings in the model may have partially adjusted for some of this bias in the measurement data on the predictions. However, this might have also contributed to the relatively small contrasts estimated between the intensity categories since we could not take into account the influence of sampling reason on the exposure levels separately.

The measurement data used to estimate quantitative exposure levels for CANJEM came from BC and Ontario, which might not necessarily be representative of exposures for jobs held predominantly in the province of Quebec. However, the sectors with the highest prevalence of exposure to wood dust in Quebec, such as construction, wood products fabrication and sawmills, tended to be the same ones identified at the national level (Labrèche *et al.*, 2013; CAREX Canada, 2015). Moreover, 99% of the BC and Ontario working population were employed in one or another of the 476 occupations in CANJEM based on data from the 2011 Canadian census (Statistics Canada, 2016a), suggesting that the information from these two provinces is compatible with the data in CANJEM. The main difference between the two sources lies in the shorter time frame covered by the measurements (23 years) compared to CANJEM. The restrictions to more recent periods in CANJEM in sensitivity analyses yielded results that were comparable to those based on the period 1930-2005. It remains, however, that the estimation of quantitative levels for jobs held prior to the late 1970s requires extrapolating the time trend while accounting for changes in technology, controls and regulations over time.

5.6 Conclusion

In this study, we modelled historical wood dust measurements to quantify semi-quantitative estimates of intensity in cells from a general-population JEM summarizing expert assessments of exposures in individual jobs. This required adapting recent approaches to combine these two sources of information while accounting for heterogeneity in the exposure profile in CANJEM cells. The development of quantitative estimates in CANJEM could be applied for the surveillance and prevention of wood dust-related diseases such as sinonasal cancers (IARC, 2012), and to estimate quantitative exposure-response relationships in etiologic studies, such as for the associations observed in previous case-control studies with lung and colorectal cancer (Siemiatycki *et al.*, 1986; Vallières *et al.*, 2015). Moreover, the framework presented here represents a starting point to develop quantitative exposure estimates to other agents present in both CANJEM and CWED, such as silica and formaldehyde with over 5000 measurements each, which should also provide further insights on the associations between the exposure levels and the intensity ratings. Lastly, this framework could also be implemented by combining other large, national exposure databases with the information available in CANJEM to develop quantitative estimates for other countries.

Acknowledgements

The authors would like to thank Marie-Pierre Sylvestre for her helpful comments on this manuscript, and Cheryl Peters for providing the CWED data.

5.7 Tables and figures

Table I. Number of samples and geometric mean and geometric standard deviation of wood dust exposure levels, overall and stratified by level of selected categorical variables.

Variable	N[1]	%[2]	ND (n)[3]	GM (mg/m^3)[4]	GSD[5]
Overall	5170	100%	412	1.62	4.4
Sample type					
Area	1451	28.1%	182	1.04	3.92
Personal	3719	71.9%	230	1.93	4.33
Source database					
BCLIMS1	1669	32.3%	206	0.93	4.49
BCLIMS2	160	3.1%	7	0.66	2.92
ONMESU	3341	64.6%	199	2.22	3.96
Province					
British Columbia	1829	35.4%	213	0.90	4.35
Ontario	3341	64.6%	199	2.22	3.96
CANJEM period					
1970-1984	678	13.1%	104	1.17	5.34
1985-2005	4492	86.9%	308	1.69	4.21
CANJEM cell rating[6]					
Low	615	11.9%	56	1.47	3.96
Medium	1232	23.8%	143	1.18	4.13
High	3324	64.3%	213	1.85	4.42

1. Number of samples
2. Relative percentage of samples by level
3. Number of non-detected samples
4. Geometric mean, computed with the robust Regression on order statistics (ROS) method for non-detects (Helsel, 2012)
5. Geometric standard deviation, computed with the robust ROS method
6. Categorical rating with the largest number of jobs in cell. Multiple imputation was used to account for ties; the figures reported are the median values of 12 iterations. Totals may therefore differ from the overall figures.

Table II. Relative effects of selected variables on wood dust exposure levels

Variable/category	RIE (%) (90% CI)[1]
Sample duration (50% increase)[2]	74.5 (73.3-75.9)
Sample year[3]	95.0 (94.2-95.7)
Intensity ratings of cell	
All jobs at low intensity	100 (reference)[4]
All jobs at medium intensity	132.4 (102.8-213.1)
All jobs at high intensity	227.0 (145.4-380.8)
Sample type	
Area	100 (reference)
Personal	150.4 (141.1-160.3)
Source database	
BC LIMS 1	100 (reference)
BC LIMS 2	143.0 (119.3-171.4)
Ontario MESU	158.2 (146.4-171.0)

1. Relative index of exposure and 90% credible interval
2. Corresponds to the effect of an increase of 50% in sampling duration. For example, using a reference duration of 60 minutes (taken as 100%), the exposure level for an increase in duration of 50% (ie., 90 minutes) is 74.5% of the reference level (ie., 25.5% lower)
3. Corresponds to the effect of an increase of one year. Using 1989 as a reference year (100%), the relative level for 1990 would be 95.0% of the reference (ie., 5.0% lower)
4. RIE of the reference level is taken as 100%

Table III. Ten occupations with the highest predicted 8-hour GM wood dust concentration for year 1989

4-digits NOC occupation code and title	N[2]	Distribution of intensity in CANJEM cells, period 1930-2005[1] Low intensity (%)	Medium intensity (%)	High intensity (%)	Predicted wood dust GM (90%CI) for year 1989 (mg/m^3) CANJEM cell ratings and occupation mean[3]	CANJEM cell ratings only[4]
7272: Cabinetmakers	279	3.9	36.3	59.8	1.67 (1.48-1.89)[5]	1.37 (1.11-1.71)
8421: Chain saw and skidder operators	14	52.1	46.7	1.2	1.64 (1.09-2.54)	0.87 (0.73-1.03)
9534: Furniture finishers and refinishers	41	14.3	57.1	28.6	1.63 (1.23-2.17)	1.13 (0.97-1.33)
7232: Tool and die makers	13	40.0	20.0	40.0	1.63 (1.10-2.50)	1.10 (0.94-1.30)
9227: Supervisors, other products manufacturing and assembly	10	0.0	0.0	100.0	1.54 (0.98-2.42)	1.68 (1.21-2.49)
9437: Woodworking machine operators	1616	5.6	24.1	70.4	1.50 (1.40-1.60)	1.43 (1.13-1.86)
9619: Other labourers in processing, manufacturing and utilities	97	38.5	26.9	34.6	1.43 (1.18-1.73)	1.08 (0.93-1.25)
7271: Carpenters	100	15.0	78.8	6.3	1.36 (1.11-1.66)	1.00 (0.84-1.22)
9224: Supervisors, furniture and fixtures manufacturing	14	23.8	38.1	38.1	1.34 (0.91-2.00)	1.15 (0.99-1.34)
9532: Furniture and fixture assemblers and inspectors	143	32.4	32.4	35.3	1.32 (1.13-1.55)	1.10 (0.95-1.28)

1. Relative percentage of jobs by intensity rating in CANJEM cells
2. Number of samples by occupation
3. Predicted geometric mean of wood dust concentrations for year 1989, sampling duration of 480 minutes, personal sample type, relative proportion of measurements in each source database, relative proportion of jobs by intensity rating of occupation, and occupation random effect mean

4. Predicted geometric mean of wood dust concentrations for year 1989, sampling duration of 480 minutes, personal sample type, relative proportion of measurements in each source database and relative proportion of jobs by intensity rating of occupation
5. Estimate (90% credible interval)

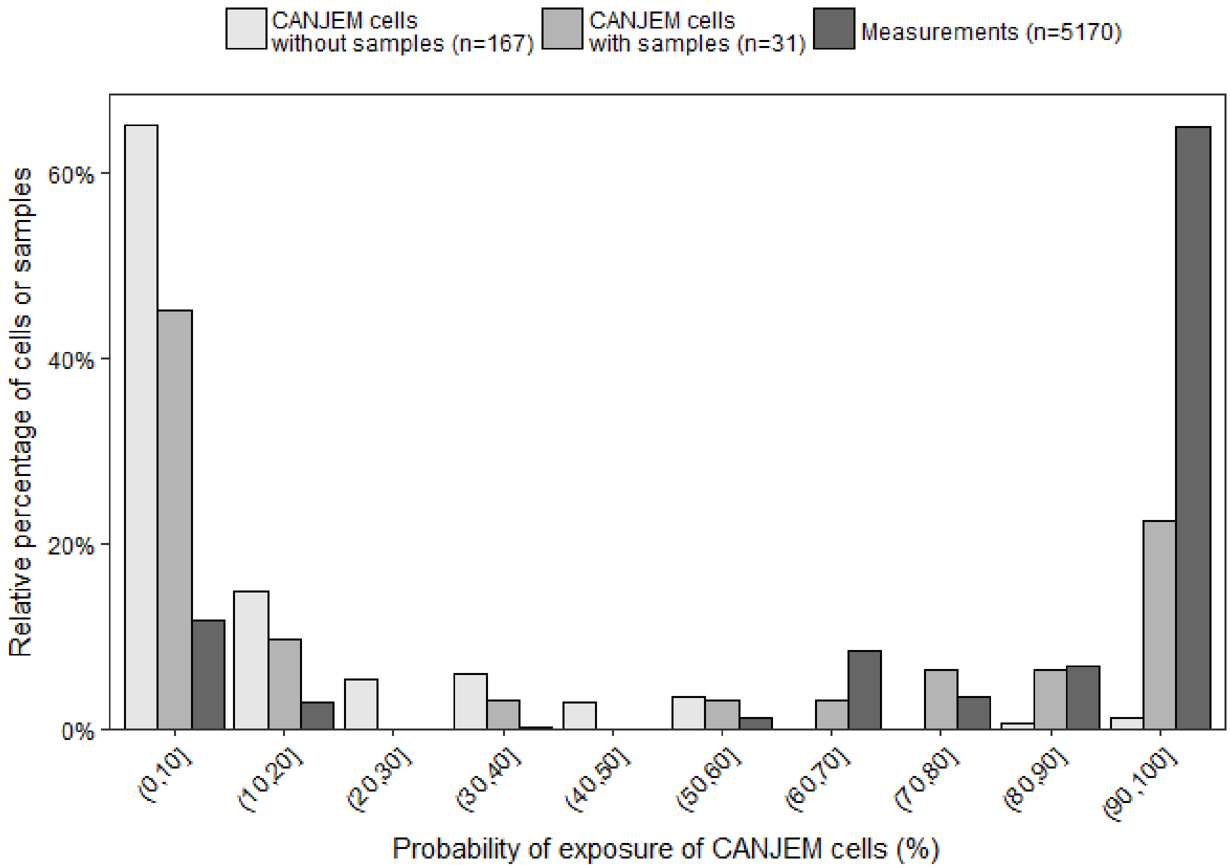
Table IV. RIEs for a theoretical cell with all jobs at medium or high intensity, relative to a cell with all jobs at low intensity, in sensitivity analyses

Analysis	Medium intensity RIE% (90% CI)[1]	High intensity RIE% (90% CI)[2]
Main analysis[3]	**132.4 (102.8-213.1)**	**227.0 (145.4-380.8)**
No constraint on intensity categories		
Proportion of jobs by rating in cell	118.0 (66.6-210.3)	204.6 (122.1-351.4)
Most frequent rating in cell[4]		
Most frequent rating, with constraint	115.4 (101.4-148.9)	164.7 (124.8-230.4)
Most frequent rating, no constraint	106.0 (78.2-143.9)	150.6 (105.9-213.0)
Alternative minimum sample size and probability of exposure by occupation		
Minimum of 5 samples/occupation and probability >0%	114.1 (101.1-154.6)	202.9 (132.8-342.1)
Minimum of 1 sample/occupation and probability >0%	119.5 (101.7-166.7)	197.7 (130.6-322.8)
Minimum of 10 samples/occupation and probability ≥5%	263.0 (137.4-506.9)	368.4 (205.8-665.2)
Minimum of 1 sample/occupation and probability ≥5%	182.6 (113.5-303.5)	269.2 (169.0-431.8)
Specific CANJEM period		
Period 1970-2005[5]	147.2 (107.4-214.1)	308.7 (187.4-527.9)
Period 1985-2005[6]	132.5 (103.0-225.4)	227.2 (147.2-431.8)

1. Relative index of exposure and 90% credible interval for all jobs at medium intensity, relative to a reference of all jobs at low intensity (taken as 100%)
2. Relative index of exposure and 90% credible interval for all jobs at high intensity, relative to a reference of all jobs at low intensity (taken as 100%)
3. Main analysis with the CANJEM ratings based on the relative proportions of jobs at medium and at high intensity in the cell in period 1930-2005, with a constraint on the order of the coefficients for the intensity ratings, and occupations entered as random effects. Restricted to occupations with at least 10 samples and at least 1 exposed job in CANJEM (probability >0%)

4. The intensity category with the highest proportion of jobs in the cell was included in the model instead of the relative proportions of jobs at medium and high intensity.
5. Based on 4029 samples from 29 occupations with at least one exposed job in period 1970-2005
6. Based on 3114 samples from 19 occupations with at least one exposed job in period 1985-2005

Figure 1. Distribution of the 198 CANJEM cells with at least one exposed job by probability of exposure, stratified by the availability of measurements (minimum of 10 samples per cell), and distribution of the measurements by the probability of exposure of the occupation in the 31 CANJEM cells with samples



5.8 References

Burns DK, Beaumont PL. (1989) The HSE National Exposure Database--(NEDB). Ann Occup Hyg; 33: 1-14.

CAREX Canada. (2015) Wood Dust. http://www.carexcanada.ca/en/wood_dust/ (Accessed February 09, 2017).

Demers P, Peters C, Davies H *et al.* (2014) Incorporating more detailed exposure assessment with quantitative estimates is assessing the burden of occupational cancer. Occup Environ Med; 71: A51-A51.

Févotte J, Dananché B, Delabre L *et al.* (2011) Matgene: A Program to Develop Job-Exposure Matrices in the General Population in France. Ann Occup Hyg; 55: 865-78.

Friesen MC, Coble JB, Lu W *et al.* (2012) Combining a job-exposure matrix with exposure measurements to assess occupational exposure to benzene in a population cohort in shanghai, china. Ann Occup Hyg; 56: 80-91.

Galea KS, Van Tongeren M, Sleeuwenhoek AJ *et al.* (2009) Trends in wood dust inhalation exposure in the UK, 1985-2005. Ann Occup Hyg; 53: 657-67.

Gérin M, Siemiatycki J, Kemper H *et al.* (1985) Obtaining occupational exposure histories in epidemiologic case-control studies. J Occup Med; 27: 420-6.

Hall A, Teschke K, Davies H *et al.* (2002) Exposure Levels and Determinants of Softwood Dust Exposures in BC Lumber Mills, 1981–1997. AIHA Journal; 63: 709-14.

Hall AL, Peters CE, Demers PA *et al.* (2014) Exposed! Or not? The diminishing record of workplace exposure in Canada. Can J Public Health; 105: e214-7.

Helsel DR. (2012) Statistics for censored environmental data using Minitab and R, 2nd edition. Hoboken, NJ: John Wiley & Sons.

Hoar SK, Morrison AS, Cole P *et al.* (1980) An occupation and exposure linkage system for the study of occupational carcinogenesis. J Occup Med; 22: 722-6.

Hon CY, Peters CE, Jardine KJ *et al.* (2017) Historical occupational isocyanate exposure levels in two Canadian provinces. J Occup Environ Hyg; 14: 1-8.

IARC. (2012) A review of human carcinogens. Part C: Arsenic, metals, fibres, and dusts. Lyon, France: International Agency for Research on Cancer.

Kauppinen T, Toikkanen J, Pukkala E. (1998) From cross-tabulations to multipurpose exposure information systems: a new job-exposure matrix. Am J Ind Med; 33: 409-17.

Kendzia B, Behrens T, Jockel KH *et al.* (2013) Welding and lung cancer in a pooled analysis of case-control studies. Am J Epidemiol; 178: 1513-25.

Kendzia B, Pesch B, Koppisch D *et al.* (2017) Modelling of occupational exposure to inhalable nickel compounds. J Expo Sci Environ Epidemiol.

Koh D-H, Bhatti P, Coble JB *et al.* (2014) Calibrating a population-based job-exposure matrix using inspection measurements to estimate historical occupational exposure to lead for a population-based cohort in Shanghai, China. J Expo Sci Environ Epidemiol; 24: 9-16.

Koppisch D, Schinkel J, Gabriel S *et al.* (2012) Use of the MEGA exposure database for the validation of the Stoffenmanager model. Ann Occup Hyg; 56: 426-39.

Labrèche F, Duguay P, Ostiguy C *et al.* (2013) Estimating occupational exposure to carcinogens in Quebec. Am J Ind Med; 56: 1040-50.

Lavoué J, Friesen MC, Burstyn I. (2013) Workplace measurements by the US Occupational Safety and Health Administration since 1979: descriptive analysis and potential uses for exposure assessment. Ann Occup Hyg; 57: 77-97.

Lavoué J, Pintos J, Van Tongeren M *et al.* (2012) Comparison of exposure estimates in the Finnish job-exposure matrix FINJEM with a JEM derived from expert assessments performed in Montreal. Occup Environ Med; 69: 465-71.

Lavoué J, Vincent R, Gérin M. (2006) Statistical modelling of formaldehyde occupational exposure levels in French industries, 1986-2003. Ann Occup Hyg; 50: 305-21.

Lippmann M. (2011) Summary and Commentary: Research Needs: Retrospective Exposure Assessment for Occupational Epidemiology. Appl Occup Environ Hyg; 6: 550-52.

Lubek BM. (2011) Province of Ontario Medical Surveillance Database. Appl Occup Environ Hyg; 10: 251-56.

Mater G, Paris C, Lavoué J. (2016) Descriptive analysis and comparison of two French occupational exposure databases: COLCHIC and SCOLA. Am J Ind Med; 59: 379-91.

Olsen E, Laursen B, Vinzents PS. (1991) Bias and random errors in historical data of exposure to organic solvents. Am Ind Hyg Assoc J; 52: 204-11.

Peters CE, Ge CB, Hall AL *et al.* (2015) CAREX Canada: an enhanced model for assessing occupational carcinogen exposure. Occup Environ Med; 72: 64-71.

Peters S, Vermeulen R, Portengen L *et al.* (2011) Modelling of occupational respirable crystalline silica exposure for quantitative exposure assessment in community-based case-control studies. J Environ Monit; 13: 3262-8.

Peters S, Vermeulen R, Portengen L *et al.* (2016) SYN-JEM: A Quantitative Job-Exposure Matrix for Five Lung Carcinogens. Ann Occup Hyg; 60: 795-811.

Plummer M. (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd international workshop on distributed statistical computing.

Sarazin P, Burstyn I, Kincl L *et al.* (2016) Trends in OSHA Compliance Monitoring Data 1979–2011: Statistical Modeling of Ancillary Information across 77 Chemicals. Ann Occup Hyg; 60: 432-52.

Sauvé J-F, Labrèche F, Richardson L *et al.* (2017) CANJEM: a general population job exposure matrix based on past expert assessments of exposure to over 250 agents. Manuscript in preparation

Scarselli A, Binazzi A, Ferrante P *et al.* (2008) Occupational exposure levels to wood dust in Italy, 1996-2006. Occup Environ Med; 65: 567-74.

Siemiatycki J, Richardson L, Gérin M *et al.* (1986) Associations Between Several Sites of Cancer and Nine Organic Dusts: Results from an Hypothesis-Generating Case-Control Study in Montreal, 1979–1983. Am J Epidemiol; 123: 235-49.

Statistics Canada. (2003) North American Industry Classification System, Canada 2002. Ottawa: Statistics Canada.

Statistics Canada. (2007) National Occupational Classification for Statistics (NOC-S). Ottawa: Statistics Canada.

Statistics Canada. (2012) National occupational classification, 2011. Ottawa: Minister of Industry.

Statistics Canada. (2016a) 2011 National Household Survey, Statistics Canada Catalogue no. 99-012-X2011033. Ottawa, ON.

Statistics Canada. (2016b) Concordance Tables for NOC 2011. Ottawa: Statistics Canada. http://noc.esdc.gc.ca/English/noc/Concordance.aspx?ver=11 (Accessed February 21, 2017).

Stewart PA, Rice C. (1990) A Source of Exposure Data for Occupational Epidemiology Studies. Appl Occup Environ Hyg; 5: 359 - 63.

Teschke K, Marion SA, Vaughan TL *et al.* (1999) Exposures to wood dust in U.S. industries and occupations, 1979 to 1997. Am J Ind Med; 35: 581-9.

Vallières E, Pintos J, Parent ME *et al.* (2015) Occupational exposure to wood dust and risk of lung cancer in two population-based case-control studies in Montreal, Canada. Environmental Health; 14: 1.

Wild P, Sauleau EA, Bourgkard E *et al.* (2002) Combining expert ratings and exposure measurements: a random effect paradigm. Ann Occup Hyg; 46: 479-87.

5.9 Appendices

Supplementary table I. Number of samples and geometric mean and geometric standard deviation of wood dust exposure levels by occupation (n=31), and the corresponding probability and intensity of exposure in CANJEM cells for period 1930-2005

	Wood dust measurement data				CANJEM cell estimates, period 1930-2005						
								Relative percentage of exposed jobs by rating			
4-digits NOC 2011 occupation code and title	N[1]	ND[2] (n)	GM[3] (mg/m3)	GSD[4]	N[5]	Nexp[6]	P[7] (%)	Low (%)	Med (%)	High (%)	Median FWI[8]
1241: Administrative assistants	13	0	3.53	6.9	663	2	0.3	100.0	0.0	0.0	0.6
2252: Industrial designers	23	3	0.75	3.0	21	4	19.0	75.0	25.0	0.0	0.8
4021: College and other vocational instructors	88	13	1.26	3.8	142	9	6.3	11.1	66.7	22.2	1.5
4031: Secondary school teachers	163	17	2.24	3.2	257	6	2.3	83.3	16.7	0.0	0.3
6733: Janitors, caretakers and building superintendents	115	7	1.66	4.4	347	55	15.9	34.5	58.2	7.3	0.6
7231: Machinists and machining and tooling inspectors	17	4	0.67	5.0	195	2	1.0	0.0	100.0	0.0	0.4
7232: Tool and die makers	13	0	2.46	3.5	64	5	7.8	40.0	20.0	40.0	0.5
7242: Industrial electricians	11	0	0.42	2.9	76	8	10.5	62.5	37.5	0.0	0.1
7271: Carpenters	100	14	2.47	3.5	343	320	93.3	15.0	78.8	6.3	5.0
7272: Cabinetmakers	279	12	2.88	3.9	105	102	97.1	3.9	36.3	59.8	15.0
7311: Construction millwrights and industrial mechanics	36	2	0.54	2.6	169	10	5.9	40.0	60.0	0.0	0.7
7321: Automotive service technicians, truck and bus mechanics and mechanical repairers	14	1	4.62	3.9	295	2	0.7	100.0	0.0	0.0	0.1
7384: Other trades and related occupations, n.e.c.	10	2	0.63	1.9	13	4	30.8	50.0	50.0	0.0	2.6
7452: Material handlers	92	14	0.92	3.4	743	56	7.5	66.1	28.6	5.4	1.0
7521: Heavy equipment operators (except crane)	18	4	0.33	3.6	119	3	2.5	33.3	66.7	0.0	0.3
7611: Construction trades helpers and labourers	63	7	1.10	3.0	488	288	59.0	50.0	48.6	1.4	0.6

4-digits NOC 2011 occupation code and title	Wood dust measurement data				CANJEM cell estimates, period 1930-2005						
								Relative percentage of exposed jobs by rating			
	N[1]	ND[2] (n)	GM[3] (mg/m3)	GSD[4]	N[5]	Nexp[6]	P[7] (%)	Low (%)	Med (%)	High (%)	Median FWI[8]
7612: Other trades helpers and labourers	10	0	2.54	1.4	31	2	6.5	0.0	100.0	0.0	0.4
8421: Chain saw and skidder operators	14	0	4.59	4.4	272	259	95.2	52.1	46.7	1.2	1.0
9224: Supervisors, furniture and fixtures manufacturing	14	0	3.11	2.8	24	21	87.5	23.8	38.1	38.1	3.0
9227: Supervisors, other products manufacturing and assembly	10	0	2.27	3.6	16	1	6.3	0.0	0.0	100.0	25.0
9431: Sawmill machine operators	1127	127	0.91	4.0	10	10	100.0	10.0	10.0	80.0	25.0
9434: Other wood processing machine operators	441	43	1.05	3.5	13	8	61.5	37.5	50.0	12.5	4.0
9436: Lumber graders and other wood processing inspectors and graders	93	10	0.60	2.6	11	10	90.9	60.0	20.0	20.0	1.0
9437: Woodworking machine operators	1616	56	2.66	4.0	57	54	94.7	5.6	24.1	70.4	25.0
9532: Furniture and fixture assemblers and inspectors	143	2	3.08	4.0	45	34	75.6	32.4	32.4	35.3	5.0
9533: Other wood products assemblers and inspectors	135	16	1.96	5.6	19	19	100.0	5.3	42.1	52.6	8.8
9534: Furniture finishers and refinishers	41	3	3.39	3.0	9	7	77.8	14.3	57.1	28.6	0.6
9536: Industrial painters, coaters and metal finishing process operators	18	2	2.12	2.5	107	8	7.5	25.0	75.0	0.0	0.6
9612: Labourers in metal fabrication	20	2	2.39	3.8	162	6	3.7	66.7	16.7	16.7	0.5
9614: Labourers in wood, pulp and paper processing	336	50	0.97	4.8	63	54	85.7	22.2	50.0	27.8	5.0
9619: Other labourers in processing, manufacturing and utilities	97	1	2.84	3.6	425	26	6.1	38.5	26.9	34.6	1.6

1. Number of samples
2. Number of non-detected samples
3. Geometric mean, computed with the robust Regression on order statistics (ROS) method for non-detects (Helsel, 2012)

4. Geometric standard deviation, computed with the robust ROS method
5. Total number of jobs evaluated
6. Number of jobs exposed to wood dust
7. Probability of exposure to wood dust
8. Median frequency-weighted intensity value among exposed jobs

Supplementary table II. Fixed and random-effects model parameters and variance components

Model parameters	Estimate[1]	SD[2]
Fixed-effects terms		
Intercept	4.033	0.221
Sample year (reference = 1978)	-0.052	0.005
Sample duration (ln[minutes])	-0.725	0.026
Sample type (reference=area)		
Personal	0.408	0.039
Source database (reference = BC LIMS1)		
BCLIMS2	0.357	0.110
Ontario MESU	0.459	0.047
Proportion of jobs in cell, medium intensity	0.280	0.229
Proportion of jobs in cell, high intensity	0.820	0.291
Variance components		
Within occupation	1.137	0.012
Between occupation	0.364	0.071
Random-effects (NOC code and title)		
1241: Administrative assistants	0.251	0.260
2252: Industrial designers	-0.493	0.232
4021: College and other vocational instructors	-0.285	0.151
4031: Secondary school teachers	0.176	0.153
6733: Janitors, caretakers and building superintendents	0.201	0.134
7231: Machinists and machining and tooling inspectors	-0.224	0.256
7232: Tool and die makers	0.392	0.252
7237: Welders and related machine operators	-0.316	0.263
7242: Industrial electricians	0.299	0.152
7271: Carpenters	0.200	0.142
7272: Cabinetmakers	-0.081	0.190
7311: Construction millwrights and industrial mechanics	0.387	0.268
7321: Automotive service technicians, truck and bus mechanics and mechanical repairers	-0.050	0.275
7384: Other trades and related occupations, n.e.c.[3]	-0.105	0.153
7452: Material handlers	-0.318	0.247
7521: Heavy equipment operators (except crane)	-0.243	0.164
7611: Construction trades helpers and labourers	0.153	0.268
7612: Other trades helpers and labourers	0.634	0.261
8421: Chain saw and skidder operators	0.153	0.241
9224: Supervisors, furniture and fixtures manufacturing	-0.097	0.283
9227: Supervisors, other products manufacturing and assembly	-0.472	0.173
9431: Sawmill machine operators	-0.240	0.098

Model parameters	Estimate[1]	SD[2]
9434: Other wood processing machine operators	-0.579	0.148
9436: Lumber graders and other wood processing inspectors and graders	0.045	0.151
9437: Woodworking machine operators	0.181	0.122
9532: Furniture and fixture assemblers and inspectors	-0.008	0.149
9533: Other wood products assemblers and inspectors	0.368	0.181
9534: Furniture finishers and refinishers	-0.147	0.235
9536: Industrial painters, coaters and metal finishing process operators	0.151	0.226
9612: Labourers in metal fabrication	-0.268	0.105
9614: Labourers in wood, pulp and paper processing	0.282	0.136

1. Median of the posterior distribution taken as the point estimate
2. Standard distribution of the posterior distribution of the model parameter
3. n.e.c.: Not elsewhere classified

Supplementary table III. Predicted geometric mean wood dust concentrations for 1989 for all occupations with at least one exposed job in CANJEM for period 1930-2005 (n=198)

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m³)[4]	
				Relative percentage of exposed jobs by rating				
	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
0015: Senior managers - trade, broadcasting and other services, n.e.c.	5	1	20.0	0.0	100.0	0.0	1.02 (0.81-1.31)	
0114: Other administrative services managers	347	4	1.2	75.0	25.0	0.0	0.81 (0.65-0.98)	
0601: Corporate sales managers	66	1	1.5	0.0	100.0	0.0	1.02 (0.81-1.31)	
0621: Retail and wholesale trade managers	629	15	2.4	80.0	20.0	0.0	0.80 (0.63-0.97)	
0631: Restaurant and food service managers	157	1	0.6	100.0	0.0	0.0	0.75 (0.56-0.94)	
0632: Accommodation service managers	39	1	2.6	0.0	100.0	0.0	1.02 (0.81-1.31)	
0651: Managers in customer and personal services, n.e.c.	17	1	5.9	0.0	100.0	0.0	1.02 (0.81-1.31)	
0711: Construction managers	89	37	41.6	81.1	18.9	0.0	0.79 (0.63-0.97)	
0731: Managers in transportation	18	1	5.6	0.0	100.0	0.0	1.02 (0.81-1.31)	
0821: Managers in agriculture	110	23	20.9	34.8	65.2	0.0	0.92 (0.77-1.09)	
0911: Manufacturing managers	145	5	3.4	40.0	40.0	20.0	1.00 (0.87-1.14)	
0912: Utilities managers	1	1	100.0	100.0	0.0	0.0	0.75 (0.56-0.94)	
1121: Human resources professionals	25	1	4.0	100.0	0.0	0.0	0.75 (0.56-0.94)	
1122: Professional occupations in business management consulting	52	1	1.9	100.0	0.0	0.0	0.75 (0.56-0.94)	
1211: Supervisors, general office and administrative support workers	55	2	3.6	100.0	0.0	0.0	0.75 (0.56-0.94)	
1215: Supervisors, supply chain, tracking and scheduling co-ordination occupations	97	8	8.2	75.0	25.0	0.0	0.81 (0.65-0.98)	

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m^3)[4]	
				Relative percentage of exposed jobs by rating				
	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
1224: Property administrators	42	6	14.3	33.3	66.7	0.0	0.92 (0.77-1.09)	
1225: Purchasing agents and officers	31	1	3.2	100.0	0.0	0.0	0.75 (0.56-0.94)	
1241: Administrative assistants	663	2	0.3	100.0	0.0	0.0	0.75 (0.56-0.94)	0.96 (0.63-1.46)
1411: General office support workers	790	4	0.5	100.0	0.0	0.0	0.75 (0.56-0.94)	
1414: Receptionists	316	1	0.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
1423: Desktop publishing operators and related occupations	21	1	4.8	0.0	100.0	0.0	1.02 (0.81-1.31)	
1431: Accounting and related clerks	551	6	1.1	100.0	0.0	0.0	0.75 (0.56-0.94)	
1432: Payroll clerks	48	1	2.1	100.0	0.0	0.0	0.75 (0.56-0.94)	
1512: Letter carriers	60	1	1.7	100.0	0.0	0.0	0.75 (0.56-0.94)	
1513: Couriers, messengers and door-to-door distributors	185	4	2.2	100.0	0.0	0.0	0.75 (0.56-0.94)	
1521: Shippers and receivers	304	11	3.6	81.8	9.1	9.1	0.83 (0.66-1.00)	
1522: Storekeepers and partspersons	188	7	3.7	57.1	28.6	14.3	0.92 (0.78-1.07)	
1523: Production logistics co-ordinators	70	2	2.9	50.0	50.0	0.0	0.87 (0.73-1.03)	
2122: Forestry professionals	3	1	33.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
2123: Agricultural representatives, consultants and specialists	6	1	16.7	0.0	100.0	0.0	1.02 (0.81-1.31)	
2131: Civil engineers	74	3	4.1	100.0	0.0	0.0	0.75 (0.56-0.94)	
2141: Industrial and manufacturing engineers	43	1	2.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
2151: Architects	25	1	4.0	100.0	0.0	0.0	0.75 (0.56-0.94)	
2154: Land surveyors	40	2	5.0	50.0	50.0	0.0	0.87 (0.73-1.03)	
2211: Chemical technologists and technicians	43	1	2.3	0.0	100.0	0.0	1.02 (0.81-1.31)	

	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m³)[4]	
				Relative percentage of exposed jobs by rating				
4-digits NOC occupation code and title	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
2223: Forestry technologists and technicians	11	5	45.5	80.0	20.0	0.0	0.80 (0.63-0.97)	
2231: Civil engineering technologists and technicians	10	2	20.0	50.0	50.0	0.0	0.87 (0.73-1.03)	
2232: Mechanical engineering technologists and technicians	28	2	7.1	0.0	100.0	0.0	1.02 (0.81-1.31)	
2233: Industrial engineering and manufacturing technologists and technicians	39	1	2.6	100.0	0.0	0.0	0.75 (0.56-0.94)	
2241: Electrical and electronics engineering technologists and technicians	58	3	5.2	100.0	0.0	0.0	0.75 (0.56-0.94)	
2242: Electronic service technicians (household and business equipment)	59	4	6.8	75.0	25.0	0.0	0.81 (0.65-0.98)	
2252: Industrial designers	21	4	19.0	75.0	25.0	0.0	0.81 (0.65-0.98)	0.49 (0.34-0.71)
2262: Engineering inspectors and regulatory officers	15	2	13.3	50.0	50.0	0.0	0.87 (0.73-1.03)	
2263: Inspectors in public and environmental health and occupational health and safety	21	1	4.8	0.0	100.0	0.0	1.02 (0.81-1.31)	
2264: Construction inspectors	19	1	5.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
4021: College and other vocational instructors	142	9	6.3	11.1	66.7	22.2	1.10 (0.95-1.31)	0.83 (0.67-1.03)
4031: Secondary school teachers	257	6	2.3	83.3	16.7	0.0	0.79 (0.62-0.96)	0.94 (0.79-1.11)
4032: Elementary school and kindergarten teachers	276	1	0.4	100.0	0.0	0.0	0.75 (0.56-0.94)	
4154: Professional occupations in religion	47	1	2.1	0.0	100.0	0.0	1.02 (0.81-1.31)	
4163: Business development officers and marketing researchers and consultants	24	1	4.2	100.0	0.0	0.0	0.75 (0.56-0.94)	
4311: Police officers (except commissioned)	112	1	0.9	0.0	100.0	0.0	1.02 (0.81-1.31)	
4313: Non-commissioned ranks of the Canadian Forces	712	1	0.1	100.0	0.0	0.0	0.75 (0.56-0.94)	
4422: Correctional service officers	21	1	4.8	0.0	100.0	0.0	1.02 (0.81-1.31)	

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m^3)[4]	
				Relative percentage of exposed jobs by rating				
	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
4423: By-law enforcement and other regulatory officers, n.e.c.	14	1	7.1	0.0	100.0	0.0	1.02 (0.81-1.31)	
5136: Painters, sculptors and other visual artists	17	3	17.6	33.3	66.7	0.0	0.92 (0.77-1.09)	
5212: Technical occupations related to museums and art galleries	2	1	50.0	0.0	100.0	0.0	1.02 (0.81-1.31)	
5223: Graphic arts technicians	6	2	33.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
5227: Support occupations in motion pictures, broadcasting, photography and the performing arts	9	1	11.1	0.0	100.0	0.0	1.02 (0.81-1.31)	
5242: Interior designers and interior decorators	9	3	33.3	33.3	33.3	33.3	1.09 (0.94-1.26)	
5243: Theatre, fashion, exhibit and other creative designers	29	5	17.2	20.0	80.0	0.0	0.96 (0.79-1.17)	
5244: Artisans and craftspersons	32	4	12.5	25.0	75.0	0.0	0.94 (0.79-1.13)	
6221: Technical sales specialists - wholesale trade	254	4	1.6	75.0	25.0	0.0	0.81 (0.65-0.98)	
6322: Cooks	306	3	1.0	66.7	33.3	0.0	0.83 (0.68-0.99)	
6331: Butchers, meat cutters and fishmongers - retail and wholesale	102	38	37.3	86.8	13.2	0.0	0.78 (0.61-0.96)	
6342: Tailors, dressmakers, furriers and milliners	204	6	2.9	50.0	50.0	0.0	0.87 (0.73-1.03)	
6343: Shoe repairers and shoemakers	147	14	9.5	50.0	42.9	7.1	0.90 (0.77-1.05)	
6344: Jewellers, jewellery and watch repairers and related occupations	66	1	1.5	100.0	0.0	0.0	0.75 (0.56-0.94)	
6345: Upholsterers	37	18	48.6	50.0	50.0	0.0	0.87 (0.73-1.03)	
6346: Funeral directors and embalmers	3	1	33.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
6411: Sales and account representatives - wholesale trade (non-technical)	408	3	0.7	33.3	66.7	0.0	0.92 (0.77-1.09)	

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m^3)[4]	
				Relative percentage of exposed jobs by rating				
	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
6421: Retail salespersons	876	15	1.7	66.7	33.3	0.0	0.83 (0.68-0.99)	
6541: Security guards and related security service occupations	383	1	0.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
6552: Other customer and information services representatives	60	2	3.3	50.0	50.0	0.0	0.87 (0.73-1.03)	
6611: Cashiers	161	1	0.6	100.0	0.0	0.0	0.75 (0.56-0.94)	
6622: Store shelf stockers, clerks and order fillers	66	6	9.1	100.0	0.0	0.0	0.75 (0.56-0.94)	
6623: Other sales related occupations	127	2	1.6	50.0	50.0	0.0	0.87 (0.73-1.03)	
6711: Food counter attendants, kitchen helpers and related support occupations	346	4	1.2	100.0	0.0	0.0	0.75 (0.56-0.94)	
6722: Operators and attendants in amusement, recreation and sport	26	1	3.8	0.0	100.0	0.0	1.02 (0.81-1.31)	
6731: Light duty cleaners	532	9	1.7	22.2	66.7	11.1	1.01 (0.86-1.19)	
6732: Specialized cleaners	89	3	3.4	100.0	0.0	0.0	0.75 (0.56-0.94)	
6733: Janitors, caretakers and building superintendents	347	55	15.9	34.5	58.2	7.3	0.95 (0.82-1.11)	1.16 (0.97-1.40)
7201: Contractors and supervisors, machining, metal forming, shaping and erecting trades and related occupations	56	3	5.4	66.7	33.3	0.0	0.83 (0.68-0.99)	
7202: Contractors and supervisors, electrical trades and telecommunications occupations	41	8	19.5	75.0	25.0	0.0	0.81 (0.65-0.98)	
7203: Contractors and supervisors, pipefitting trades	16	1	6.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
7204: Contractors and supervisors, carpentry trades	21	17	81.0	35.3	58.8	5.9	0.94 (0.81-1.10)	
7205: Contractors and supervisors, other construction trades, installers, repairers and servicers	51	19	37.3	57.9	42.1	0.0	0.85 (0.71-1.01)	

	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m³)[4]	
				Relative percentage of exposed jobs by rating				
4-digits NOC occupation code and title	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
7231: Machinists and machining and tooling inspectors	195	2	1.0	0.0	100.0	0.0	1.02 (0.81-1.31)	0.82 (0.54-1.22)
7232: Tool and die makers	64	5	7.8	40.0	20.0	40.0	1.10 (0.94-1.30)	1.63 (1.10-2.50)
7233: Sheet metal workers	112	12	10.7	41.7	58.3	0.0	0.90 (0.76-1.06)	
7235: Structural metal and platework fabricators and fitters	32	4	12.5	25.0	75.0	0.0	0.94 (0.79-1.13)	
7236: Ironworkers	46	6	13.0	66.7	33.3	0.0	0.83 (0.68-0.99)	
7237: Welders and related machine operators	260	3	1.2	100.0	0.0	0.0	0.75 (0.56-0.94)	
7241: Electricians (except industrial and power system)	124	71	57.3	66.2	33.8	0.0	0.83 (0.68-0.99)	
7242: Industrial electricians	76	8	10.5	62.5	37.5	0.0	0.84 (0.70-1.00)	0.61 (0.39-0.93)
7244: Electrical power line and cable workers	40	1	2.5	100.0	0.0	0.0	0.75 (0.56-0.94)	
7245: Telecommunications line and cable workers	12	2	16.7	100.0	0.0	0.0	0.75 (0.56-0.94)	
7246: Telecommunications installation and repair workers	38	4	10.5	100.0	0.0	0.0	0.75 (0.56-0.94)	
7247: Cable television service and maintenance technicians	1	1	100.0	100.0	0.0	0.0	0.75 (0.56-0.94)	
7251: Plumbers	101	26	25.7	88.5	11.5	0.0	0.78 (0.60-0.96)	
7252: Steamfitters, pipefitters and sprinkler system installers	97	10	10.3	70.0	30.0	0.0	0.82 (0.67-0.99)	
7271: Carpenters	343	320	93.3	15.0	78.8	6.3	1.00 (0.84-1.22)	1.36 (1.11-1.66)
7272: Cabinetmakers	105	102	97.1	3.9	36.3	59.8	1.37 (1.11-1.71)	1.67 (1.48-1.89)
7281: Bricklayers	105	11	10.5	81.8	18.2	0.0	0.79 (0.63-0.97)	
7282: Concrete finishers	39	8	20.5	62.5	37.5	0.0	0.84 (0.70-1.00)	
7283: Tilesetters	16	5	31.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
7284: Plasterers, drywall installers and finishers and lathers	47	12	25.5	58.3	41.7	0.0	0.85 (0.71-1.01)	

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m^3)[4]	
				Relative percentage of exposed jobs by rating				
	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
7291: Roofers and shinglers	38	11	28.9	81.8	18.2	0.0	0.79 (0.63-0.97)	
7292: Glaziers	11	6	54.5	50.0	50.0	0.0	0.87 (0.73-1.03)	
7293: Insulators	23	6	26.1	83.3	16.7	0.0	0.79 (0.62-0.96)	
7294: Painters and decorators (except interior decorators)	193	63	32.6	20.6	77.8	1.6	0.96 (0.80-1.17)	
7295: Floor covering installers	25	5	20.0	20.0	0.0	80.0	1.43 (1.10-1.95)	
7301: Contractors and supervisors, mechanic trades	84	2	2.4	50.0	50.0	0.0	0.87 (0.73-1.03)	
7305: Supervisors, motor transport and other ground transit operators	25	1	4.0	0.0	100.0	0.0	1.02 (0.81-1.31)	
7311: Construction millwrights and industrial mechanics	169	10	5.9	40.0	60.0	0.0	0.90 (0.76-1.06)	0.83 (0.62-1.11)
7312: Heavy-duty equipment mechanics	61	4	6.6	25.0	50.0	25.0	1.07 (0.93-1.24)	
7313: Refrigeration and air conditioning mechanics	19	2	10.5	50.0	50.0	0.0	0.87 (0.73-1.03)	
7314: Railway carmen/women	66	22	33.3	22.7	54.5	22.7	1.07 (0.93-1.24)	
7315: Aircraft mechanics and aircraft inspectors	115	2	1.7	50.0	50.0	0.0	0.87 (0.73-1.03)	
7316: Machine fitters	30	2	6.7	50.0	50.0	0.0	0.87 (0.73-1.03)	
7321: Automotive service technicians, truck and bus mechanics and mechanical repairers	295	2	0.7	100.0	0.0	0.0	0.75 (0.56-0.94)	1.10 (0.73-1.70)
7322: Motor vehicle body repairers	97	3	3.1	66.7	33.3	0.0	0.83 (0.68-0.99)	
7331: Oil and solid fuel heating mechanics	11	3	27.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
7332: Appliance servicers and repairers	19	1	5.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
7333: Electrical mechanics	19	1	5.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
7371: Crane operators	77	3	3.9	100.0	0.0	0.0	0.75 (0.56-0.94)	

	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m3)[4]	
				Relative percentage of exposed jobs by rating				
4-digits NOC occupation code and title	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
7372: Drillers and blasters - surface mining, quarrying and construction	45	1	2.2	100.0	0.0	0.0	0.75 (0.56-0.94)	
7381: Printing press operators	106	2	1.9	0.0	50.0	50.0	1.32 (1.09-1.62)	
7384: Other trades and related occupations, n.e.c.	13	4	30.8	50.0	50.0	0.0	0.87 (0.73-1.03)	0.83 (0.52-1.30)
7441: Residential and commercial installers and servicers	51	29	56.9	24.1	27.6	48.3	1.21 (1.02-1.45)	
7445: Other repairers and servicers	25	1	4.0	0.0	100.0	0.0	1.02 (0.81-1.31)	
7451: Longshore workers	87	14	16.1	21.4	71.4	7.1	0.99 (0.84-1.18)	
7452: Material handlers	743	56	7.5	66.1	28.6	5.4	0.85 (0.71-1.01)	0.77 (0.62-0.94)
7511: Transport truck drivers	784	54	6.9	63.0	31.5	5.6	0.86 (0.72-1.02)	
7514: Delivery and courier service drivers	499	12	2.4	66.7	33.3	0.0	0.83 (0.68-0.99)	
7521: Heavy equipment operators (except crane)	119	3	2.5	33.3	66.7	0.0	0.92 (0.77-1.09)	0.67 (0.44-0.99)
7522: Public works maintenance equipment operators and related workers	36	1	2.8	100.0	0.0	0.0	0.75 (0.56-0.94)	
7531: Railway yard and track maintenance workers	60	4	6.7	100.0	0.0	0.0	0.75 (0.56-0.94)	
7532: Water transport deck and engine room crew	114	3	2.6	33.3	66.7	0.0	0.92 (0.77-1.09)	
7611: Construction trades helpers and labourers	488	288	59.0	50.0	48.6	1.4	0.88 (0.74-1.03)	0.69 (0.54-0.87)
7612: Other trades helpers and labourers	31	2	6.5	0.0	100.0	0.0	1.02 (0.81-1.31)	1.19 (0.77-1.86)
7621: Public works and maintenance labourers	53	7	13.2	28.6	71.4	0.0	0.93 (0.78-1.12)	
7622: Railway and motor transport labourers	17	1	5.9	100.0	0.0	0.0	0.75 (0.56-0.94)	
8211: Supervisors, logging and forestry	5	3	60.0	66.7	33.3	0.0	0.83 (0.68-0.99)	
8255: Contractors and supervisors, landscaping, grounds maintenance and horticulture services	15	1	6.7	0.0	100.0	0.0	1.02 (0.81-1.31)	

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m³)[4]	
				Relative percentage of exposed jobs by rating				
	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
8421: Chain saw and skidder operators	272	259	95.2	52.1	46.7	1.2	0.87 (0.73-1.03)	1.64 (1.09-2.54)
8422: Silviculture and forestry workers	9	4	44.4	75.0	25.0	0.0	0.81 (0.65-0.98)	
8431: General farm workers	596	77	12.9	53.2	42.9	3.9	0.88 (0.75-1.03)	
8432: Nursery and greenhouse workers	13	1	7.7	0.0	100.0	0.0	1.02 (0.81-1.31)	
8612: Landscaping and grounds maintenance labourers	87	7	8.0	14.3	85.7	0.0	0.97 (0.80-1.20)	
8614: Mine labourers	65	4	6.2	25.0	75.0	0.0	0.94 (0.79-1.13)	
8616: Logging and forestry labourers	25	13	52.0	84.6	15.4	0.0	0.79 (0.62-0.96)	
9212: Supervisors, petroleum, gas and chemical processing and utilities	28	3	10.7	100.0	0.0	0.0	0.75 (0.56-0.94)	
9214: Supervisors, plastic and rubber products manufacturing	15	2	13.3	50.0	50.0	0.0	0.87 (0.73-1.03)	
9215: Supervisors, forest products processing	9	5	55.6	80.0	0.0	20.0	0.89 (0.70-1.06)	
9217: Supervisors, textile, fabric, fur and leather products processing and manufacturing	111	7	6.3	28.6	57.1	14.3	1.00 (0.87-1.16)	
9222: Supervisors, electronics manufacturing	12	1	8.3	100.0	0.0	0.0	0.75 (0.56-0.94)	
9224: Supervisors, furniture and fixtures manufacturing	24	21	87.5	23.8	38.1	38.1	1.15 (0.99-1.34)	1.34 (0.91-2.00)
9227: Supervisors, other products manufacturing and assembly	16	1	6.3	0.0	0.0	100.0	1.68 (1.21-2.49)	1.54 (0.98-2.42)
9241: Power engineers and power systems operators	107	5	4.7	40.0	40.0	20.0	1.00 (0.87-1.14)[5]	
9411: Machine operators, mineral and metal processing	37	1	2.7	0.0	100.0	0.0	1.02 (0.81-1.31)	
9412: Foundry workers	79	2	2.5	100.0	0.0	0.0	0.75 (0.56-0.94)	
9414: Concrete, clay and stone forming operators	50	1	2.0	0.0	100.0	0.0	1.02 (0.81-1.31)	

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005						Predicted wood dust GM (90%CI) for year 1989 (mg/m^3)[4]	
				Relative percentage of exposed jobs by rating				
	N[1]	Nexp[2]	P[3] (%)	Low (%)	Med (%)	High (%)	Ratings only	Combined ratings and random effect
9416: Metalworking and forging machine operators	149	2	1.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
9417: Machining tool operators	175	2	1.1	50.0	50.0	0.0	0.87 (0.73-1.03)	
9421: Chemical plant machine operators	98	5	5.1	0.0	100.0	0.0	1.02 (0.81-1.31)	
9422: Plastics processing machine operators	26	2	7.7	0.0	50.0	50.0	1.32 (1.09-1.62)	
9431: Sawmill machine operators	10	10	100.0	10.0	10.0	80.0	1.48 (1.14-2.00)	0.92 (0.85-1.00)
9432: Pulp mill machine operators	5	2	40.0	100.0	0.0	0.0	0.75 (0.56-0.94)	
9433: Papermaking and finishing machine operators	12	2	16.7	50.0	50.0	0.0	0.87 (0.73-1.03)	
9434: Other wood processing machine operators	13	8	61.5	37.5	50.0	12.5	0.97 (0.84-1.12)	0.76 (0.68-0.84)
9435: Paper converting machine operators	68	3	4.4	100.0	0.0	0.0	0.75 (0.56-0.94)	
9436: Lumber graders and other wood processing inspectors and graders	11	10	90.9	60.0	20.0	20.0	0.94 (0.79-1.09)	0.52 (0.43-0.64)
9437: Woodworking machine operators	57	54	94.7	5.6	24.1	70.4	1.43 (1.13-1.86)	1.50 (1.40-1.60)
9441: Textile fibre and yarn, hide and pelt processing machine operators and workers	143	12	8.4	41.7	50.0	8.3	0.93 (0.80-1.08)	
9445: Fabric, fur and leather cutters	117	3	2.6	66.7	33.3	0.0	0.83 (0.68-0.99)	
9446: Industrial sewing machine operators	680	1	0.1	100.0	0.0	0.0	0.75 (0.56-0.94)	
9447: Inspectors and graders, textile, fabric, fur and leather products manufacturing	41	1	2.4	100.0	0.0	0.0	0.75 (0.56-0.94)	
9461: Process control and machine operators, food, beverage and associated products processing	173	4	2.3	25.0	50.0	25.0	1.07 (0.93-1.24)	
9462: Industrial butchers and meat cutters, poultry preparers and related workers	100	21	21.0	95.2	4.8	0.0	0.76 (0.58-0.95)	
9471: Plateless printing equipment operators	20	1	5.0	100.0	0.0	0.0	0.75 (0.56-0.94)	

4-digits NOC occupation code and title	CANJEM cell estimates, period 1930-2005 N[1]	Nexp[2]	P[3] (%)	Relative percentage of exposed jobs by rating: Low (%)	Med (%)	High (%)	Predicted wood dust GM (90%CI) for year 1989 (mg/m^3)[4]: Ratings only	Combined ratings and random effect
9472: Camera, platemaking and other prepress occupations	26	1	3.8	100.0	0.0	0.0	0.75 (0.56-0.94)	
9474: Photographic and film processors	23	1	4.3	0.0	0.0	100.0	1.68 (1.21-2.49)	
9523: Electronics assemblers, fabricators, inspectors and testers	62	2	3.2	50.0	50.0	0.0	0.87 (0.73-1.03)	
9524: Assemblers and inspectors, electrical appliance, apparatus and equipment manufacturing	60	2	3.3	50.0	50.0	0.0	0.87 (0.73-1.03)	
9526: Mechanical assemblers and inspectors	22	5	22.7	40.0	60.0	0.0	0.90 (0.76-1.06)	
9532: Furniture and fixture assemblers and inspectors	45	34	75.6	32.4	32.4	35.3	1.10 (0.95-1.28)	1.32 (1.13-1.55)
9533: Other wood products assemblers and inspectors	19	19	100.0	5.3	42.1	52.6	1.31 (1.08-1.61)	1.30 (1.10-1.55)
9534: Furniture finishers and refinishers	9	7	77.8	14.3	57.1	28.6	1.13 (0.97-1.33)	1.63 (1.23-2.17)
9535: Plastic products assemblers, finishers and inspectors	16	1	6.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
9536: Industrial painters, coaters and metal finishing process operators	107	8	7.5	25.0	75.0	0.0	0.94 (0.79-1.13)	0.81 (0.55-1.19)
9537: Other products assemblers, finishers and inspectors	158	13	8.2	53.8	30.8	15.4	0.93 (0.80-1.08)	
9611: Labourers in mineral and metal processing	93	6	6.5	50.0	50.0	0.0	0.87 (0.73-1.03)	
9612: Labourers in metal fabrication	162	6	3.7	66.7	16.7	16.7	0.90 (0.75-1.06)	1.05 (0.73-1.51)
9614: Labourers in wood, pulp and paper processing	63	54	85.7	22.2	50.0	27.8	1.10 (0.95-1.27)	0.84 (0.74-0.95)
9615: Labourers in rubber and plastic products manufacturing	12	1	8.3	0.0	100.0	0.0	1.02 (0.81-1.31)	
9617: Labourers in food, beverage and associated products processing	122	2	1.6	100.0	0.0	0.0	0.75 (0.56-0.94)	
9619: Other labourers in processing, manufacturing and utilities	425	26	6.1	38.5	26.9	34.6	1.08 (0.93-1.25)	1.43 (1.18-1.73)

1. Total number of jobs evaluated
2. Number of jobs exposed to wood dust
3. Probability of exposure to wood dust
4. Predicted GM for the year 1989, a sampling duration of 480 minutes, personal sample type, relative proportion of measurements in each source database, and the relative proportion of exposed jobs by intensity rating in the cell
5. Point estimate, 90% credible interval in parentheses

Supplementary file 1. JAGS program used for the main analysis

```
# JAGS program used for the main model
#   Model defined by
# - Occupations (4-digit NOC codes) as random effects (n=31)
# - CANJEM rating as proportion of jobs
#   at medium and proportion of jobs at high
#       with constraint on low/medium/high
# - other predictors:
#     1) sample year,
#     2) sample duration,
#     3) personal sample type,
#     4) Source database = BC LIMS2,
#     5) Source database = Ontario MESU

    model{

    # A) PRIORS

    # 1) Intercept
    alpha ~ dnorm(0, 0.0001)
    # 2) Coefficients for fixed effects (except intensity)
    #   where nbeta = 5
    for(i in 1:nbeta) {beta[i] ~ dnorm(0, 0.1756)}

    # 3) Coefficients for intensity levels with constraint
    # Folded normal distributions
    for(i in 1:2){beta.C0[i] ~ dnorm(0, 0.02145)T(0,)}
    # Sort to put constraint medium < high
    beta.C[1:2] <- sort(beta.C0)

    # 4) Occupation means
    #   where n.noc = 31 occupations
    for(i in 1:n.noc){
    mu.noc[i] ~ dnorm(0, tau.noc)
    }

    # 5) Residual SD
    tau <- pow(sigma, -2)
    sigma ~ dunif(0, 10)

    # 6)Between-occupation SD
    sigma.noc ~ dunif(0, 10)
    tau.noc <- pow(sigma.noc, -2)

    # B) LIKELIHOOD
    # Loop over samples (n=5170)
    for(i in 1:n){
        # Imputation for non-detects
```

```
      isAboveLOD[i] ~ dinterval(logconc[i], loglod[i])
        # Likelihood
      logconc[i] ~ dnorm(mu[i], tau)
        mu[i] <- alpha + mu.noc[vec.noc[i]] + inprod(beta, modmat[i,]) +
                          inprod(beta.C, modmat.C[i,])
    }
# Where
# modmat = design matrix for predictors
#     other than occupation and CANJEM rating
# modmat.C = design matrix for CANJEM ratings
#   (proportion of jobs at medium and proportion at high)
# vec.noc = index of occupations

# C) PREDICTIONS
  # 1) For 100% of jobs by intensity rating, duration of 480 minutes
  for(i in 1:3){
   pred.C480[i] <- exp(alpha + inprod(beta, Cpredmat480[i,]) +
                          inprod(beta.C, Cpredmat.C[i,]))
}
# Where
    # Cpredmat = prediction matrix for
    #   predictors other than occupation and CANJEM rating
    # Cpredmat.C = prediction matrix for CANJEM rating

  # 2) For occupations, duration of 480 minutes
  for(i in 1:n.noc){
    pred.noc480[i] <- exp(alpha + mu.noc[predmat.noc[i]] +
                 inprod(beta, predmat480[i,]) +
                 inprod(beta.C, predmat.C[i,]))
  }
    # Where
        # n.noc = 31 occupations
        # predmat = prediction matrix for predictors
      #   other than occupation and CANJEM rating
        # predmat.C = prediction matrix for CANJEM rating
      #   (proportion of jobs at medium and high levels in cell)
        # predmat.noc = index of occupation code

  # 3) For all cells with P>0% in CANJEM, based on
     #      proportions of jobs at low, medium and high
        # n.nocAll = 198 occupations
        # predmatAll480 = prediction matrix for predictors other
                     than CANJEM ratings
        # predmatAll.C = prediction matrix for CANJEM rating of occupations
  for(i in 1:n.nocAll){
  pred.nocAll480[i] <- exp(alpha + inprod(beta, predmatAll480[i,]) +
                    inprod(beta.C, predmatAll.C[i,]))
  }
  }
```

Chapitre 6. A hybrid expert approach for retrospective assessment of occupational exposures in a population-based case-control study

A hybrid expert approach for retrospective assessment of occupational exposures in a population-based case-control study

Jean-François Sauvé [1,2], Jérôme Lavoué [1,2], Marie-Élise Parent [2,3,4]

1. Université de Montréal, School of Public Health, Department of environmental and occupational health, Montréal, Québec, Canada
2. Centre de recherche du CHUM, Montréal, Québec, Canada
3. INRS-Institut Armand-Frappier, Université du Québec, Laval, Québec, Canada
4. Université de Montréal, School of Public Health, Department of Social and Preventive Medicine, Montréal, Québec, Canada

Jean-François Sauvé a contribué de façon majeure à la préparation des données, à leur analyse et à l'interprétation des résultats, ainsi qu'à la rédaction et à la préparation du manuscrit.

6.1 Abstract

Background: PROtEuS is a population-based case-control study of prostate cancer in Montreal, Canada, comprising approximately 4000 subjects. Interviews collected detailed lifetime occupational histories. To facilitate and improve the retrospective occupational exposure assessment for over 300 agents, a hybrid method was designed combining expert review of jobs with job-exposure profiles (JEPs) summarizing evaluations from previous studies by occupation (n=1571).

Objectives: To describe the hybrid method and its impacts on the exposures assigned compared to a previous study using a traditional expert approach.

Methods: The PROtEuS experts evaluated 16,065 jobs to assign semi-quantitative ratings by reliability, concentration and frequency of exposure (3 categories each) to 313 agents. These were compared to jobs from a lung cancer study used as a source for JEPs for 90 blue collar occupations and 203 agents common between the two sets. Endpoints evaluated included differences in the number of exposures and in the distribution of ratings across jobs, and the variability in the exposure levels of jobs within an occupation.

Results: PROtEuS job had on average 0.3 more exposures and a higher proportion of definite reliability (61%) relative to the Lung cancer study (55%). There was a trend of lower variability in the ratings assigned by occupation and agent in PROtEuS jobs for all metrics, particularly for concentration. Use of the hybrid method resulted in a coding time under 1h/job, compared to 2h/job in previous studies.

Conclusions: The method provided increased efficiency and reliability in the assessments. While the ratings assigned were more homogeneous, significant between-job variability remained within occupations, suggesting that the experts used the job description to modulate the information in JEPs.

6.2 Introduction

Occupational exposure assessment in community-based studies involves evaluating jobs distributed across a range of industries and workplaces, often over a period of several decades, and with limited historical measurement data available. Job-exposure matrices (JEMs), expert review of individual job descriptions and self-reports have long been the only feasible approaches to estimate past exposures (Teschke *et al.*, 2002). A JEM is a cross-tabulation with a list of occupations on one axis, and a list of agents on the other. Each matrix cell provides exposure estimates specific to a given combination of occupation and agent (Kauppinen *et al.*, 1998; Kromhout and Vermeulen, 2001; Teschke, 2003). JEMs provide automatic assignment of exposure based on occupation title (or industry) and, in some cases, employment period. This results in the same estimate being assigned to all jobs sharing the same occupation, without consideration for potential heterogeneity in exposures. In contrast, expert judgment applied to individual jobs descriptions (Gérin *et al.*, 1985) can use information on tasks, processes and other information reported by subjects to assign job-specific exposures. For this reason, the expert approach has been recognized as the reference method for retrospective community-based studies (Bouyer and Hémon, 1993). However, with nowadays large study samples, the labor-intensive process for assigning exposure estimates has become prohibitive.

Refining the expert method to increase its efficiency while maintaining its ability to provide accurate job-specific estimates of exposure has been an active area of research in recent years. This includes assigning exposures by applying predefined decision rules to questionnaire responses, enabling an automatic classification of jobs as unexposed, exposed, or necessitating review (Fritschi *et al.*, 2009; Fritschi *et al.*, 2012; Pronk *et al.*, 2012). Such rules can be based on expert judgment, but they may also be informed by existing sources of exposure information such as JEMs (Behrens *et al.*, 2012) or exposure evaluations from past studies (Wheeler *et al.*, 2013; Friesen *et al.*, 2016).

In this paper, we describe an approach to retrospective exposure assessment developed for PROtEuS (Prostate Cancer & Environment Study), a population-based case-control study comprising approximately 4000 subjects in Montreal. One objective of PROtEuS is to explore potential associations between the risk of prostate cancer and occupational exposure to some 300 chemical and physical agents. Due to the large number of jobs to review (over 16,000) and limited resources to assess exposure, a hybrid method was devised in which the traditional expert method was informed by historical exposure evaluations from previous Montreal-based case-control studies, summarized into profiles by occupation ("job-exposure profiles", or JEPs). This method was developed to not only decrease the time spent on evaluating each job, but also to enrich the industrial hygiene information readily available to experts for coding, to increase inter-expert consistency, to train new experts and to lower the probability of experts missing exposures for complex jobs. However, since this method provides experts with exposure distributions for a given occupation, there could be a risk that their assignments would lean too close to the JEPs, thereby overlooking idiosyncrasies in specific job circumstances. To this end, we also conducted an evaluation of this method by comparing the exposure data underlying the backbone JEPs to the exposure data generated from the hybrid approach. Our *a priori* hypotheses were that the latter would provide higher confidence in the assessments, increase the number of exposures assigned to jobs, and lead to lower variability in exposure between jobs within the same occupation.

6.3 Methods

PROtEuS study population

The PROtEuS study population has been described in detail previously (Blanc-Lapierre *et al.*, 2015). Briefly, eligible subjects were men aged ≤75 years at diagnosis or recruitment, Canadian citizens, registered on the permanent electoral list and residing among the 39 electoral districts of the greater Montreal area. Eligible cases were all patients newly diagnosed with primary histologically confirmed prostate cancer from September 2005 to December 2009, actively ascertained from pathology departments across the main Montreal hospitals serving the French-speaking population

For each subject, questionnaires and semi-structured in-person interviews were used to collect detailed lifestyle and socio-economic factors, medical history and complete occupational history covering each job held during lifetime. A total of 4013 subjects were interviewed, with 1966 recruited as cases (49%) and 2047 (51%) as controls. Proxy respondents (<4%) completed the interview when subjects were unable to do so. Information collected for each job included job title, company name, tasks performed, products or equipment used, use of protective measures, and descriptions of the work environment. Specialized questionnaires were also used to collect additional information for occupations with a more complex exposure profile, such as welders or auto mechanics. A team of trained chemists-hygienists, blind to case/control status, then reviewed the job histories to assign standardized job and industry titles for each job. Job titles were coded in the 1971 Canadian Classification and Dictionary of Occupations (CCDO) (Department of Employment and Immigration, 1971), and the 1980 Canadian Standardized Industrial Classifications (SIC) (Statistics Canada, 1980) for industries.

Development of the job-exposure profiles for PROtEuS

Source databases of exposure assessments

The JEPs were created from the exposure data of two population-based case-control studies conducted in Montreal (Table I). The first study ("Multisite") was conducted from 1979-1986 and included men aged 35-70 years diagnosed with one of 19 different sites of cancer (n=3730) (Siemiatycki, 1991). The second study ("Lung cancer study") included men and women aged 29-75 years diagnosed with a lung malignancy between 1996-1998 (n=1231). Cases were actively recruited from all major Montreal area hospitals, were Canadian citizens and living within the Montreal region. Population controls (Multisite n=533, Lung cancer study n=1513) were randomly selected from electoral lists, frequency-matched to cases by age and sex (Lung cancer study only). The data collection methods, including occupational questionnaires, were analogous to those described for PROtEuS.

In both source studies, occupational exposures were assigned using the traditional expert-based approach described in Gérin *et al.* (1985) and elsewhere (Siemiatycki, 1991; Parent *et al.*, 2007). A team of experts reviewed each job description to assign exposures to a predefined list of approximately 300 chemical, physical and biological agents, including mixtures (e.g. plating solutions) and general categories (e.g. pesticides). A job was considered exposed to an agent if it was present in the workplace at a level above those found in the general environment. Exposure was rated in three dimensions, each using three categories: reliability, or degree of confidence in the assessment (possible, probable, definite), concentration (low, medium, high) and frequency, as the relative percentage of the working week with exposure (<5%, 5-30%, >30%). For some agents, a list of occupations and processes representative of low, medium and high concentration served as benchmarks to standardize the exposure assessment (eg. Vida *et al.*, 2010; Pintos *et al.*, 2012). Each job was evaluated separately by two experts, and the final assessment was based on a consensus. An example of expert assessment applied to diesel exhaust exposure for motor vehicle mechanic jobs is presented in Parent *et al.* (2007).

Job-exposure profiles

A JEP presents a comprehensive view of all exposures assigned to jobs for each specific occupation in the source databases. The core of the JEPs consists of descriptive tables summarizing the exposures assigned for a list of 1571 7-digit occupations that had at least one job evaluated in previous studies (Figure 1). For the Lung cancer study, only jobs held by male subjects (n=1661) were retained, with exposures covering 289 agents. These were supplemented by data from the Multisite study for 23 agents, predominantly metals with a physical form (e.g., “iron fumes”). Each summary table lists the agents with at least 1 job exposed in that occupation. For each agent, the number of exposed jobs and their distribution across the reliability, concentration and frequency categories are provided. Colors are used as visual clues to represent the variability in the distribution of the categorical ratings assigned to exposed jobs: Green when >75% of jobs were assigned to one category, yellow for 50-75% (or based on only 2 jobs), and red for <50% (or based on only 1 job). An illustration of a JEP for Combination welders (CCDO 8335-126) containing exposure data to 111 agents is presented in Supplementary file 1 (available online at expostats.ca/jeanf/chapitre6/suppl_file_1.xls). A subset of this JEP featuring 5 agents is presented in Figure 2.

JEP add-ons

Along with the descriptive summaries of the exposure data from previous studies, the JEP framework featured additional components to assist the experts:

Job title definitions: To facilitate the selection of the most appropriate JEP for a job description, the JEP interface enabled to search through the 8,000 unique occupations in the CCDO classification, and to compare their descriptions.

Agent definitions: A short definition of each agent was provided, including its chemical/physical properties, potential sources and uses, correlated exposures and, when available, relevant historical measurements from the literature.

Annotations: The JEPs contained short comments or justifications made by the experts when assigning specific exposures in previous studies. A comprehensive review was also conducted for 295 occupations with a more complex exposure profile (eg, construction labourers) to add more detailed comments based on prior knowledge or industrial hygiene literature to guide experts in adjusting exposure levels based on specific tasks and circumstances reported by subjects. For instance, asbestos exposure among auto mechanics would be associated with low concentration, except when a job entailed brake repair work with medium concentration. For simplicity, this review was performed on groups of related occupations (eg. driving occupations) rather than independently for each occupation.

Link to source data: This enabled experts to go back to the original exposure assessments of individual jobs from the source studies to compare occupational circumstances.

Application of the hybrid expert method in PROtEuS

Four experts carried out the exposure assessment in PROtEuS. Of these, two were senior experts who had not only participated in the development of the traditional expert approach back in the 1970s in Montreal, but who had also applied it in the two source studies.

The exposure assessment approach in PROtEuS followed the same principles as the source studies but differed in two aspects. First, the highest frequency category assigned in previous studies covered the range of 30% to 100% of the workweek. In PROtEuS, this category was split in two: one for >30% to 90% of the workweek, and another for >90% to better represent continuous exposure. The second difference with previous studies is the reliance on the data in the JEPs to guide the experts in their evaluations.

When evaluating exposures for a job, the experts could retrieve the relevant JEP based on the 7-digit CCDO code, and assign exposures based on the job description and the information provided in the JEP. This method thus meets halfway between JEMs, where a fixed set of exposures are assigned to jobs based on the occupation, and the traditional expert method resting on the evaluation of individual job descriptions. In this hybrid method, the experts could omit exposures suggested by the JEP and/or assign additional exposures not listed in the JEP. They were also not constrained to a single JEP and could draw information from other occupations when relevant to a given job description. For example, a janitor reporting regular plumbing tasks could entail the use of JEPs from plumbing occupations to cover a wider spectrum of exposures. When evaluating exposures for a job with no JEP available, the experts could also use data from related occupations for guidance.

Comparison of the exposures assigned in PROtEuS to those from a source study coded using the traditional expert method

The hybrid approach provided experts with additional information and guidance to facilitate their assessments. On the other hand, experts might have tended to assign exposures that overly resembled the JEPs, resulting in reduced variability in exposures. We therefore conducted analyses to examine if the experts tended to assign more exposures to jobs and at a higher reliability with the hybrid method compared to the traditional approach, and if the use of the hybrid method resulted in lower variability in the exposure ratings between jobs within an occupation. The exposures assigned in PROtEuS were compared to those in the Lung cancer study serving as a reference. Some comparisons were based on the individual jobs, and some on jobs summarized by combination of occupation and agent.

Since the Lung cancer study data was used to inform the experts in PROtEuS and contained a different set of job histories, this comparison does not represent an evaluation of the validity of the hybrid method. The latter would have required comparing exposures assigned independently with both methods over a common sample of job descriptions.

Data selection

The comparisons were restricted to blue collar occupations for which at least 10 jobs were evaluated in both PROtEuS and the Lung cancer study (jobs held by male subjects only). Blue-collar occupations were those in 4-digit CCDO unit groups classified as skilled, semiskilled, unskilled and farming occupations in the Pineo-Porter-McRoberts socioeconomic classification (Pineo *et al.*, 1977; Aronson *et al.*, 2000). We retained the exposure data associated with 90 occupations (Supplementary table I) encompassing 4318 jobs in PROtEuS and 3022 in the Lung cancer study (Figure 1). White collar occupations were excluded as there was relatively fewer jobs in the Lung cancer study (22% of all jobs), fewer JEPs (n=447) and were generally exposed to fewer agents. Lastly, we retained 203 agents (listed in Supplementary table II) evaluated in both studies and had at least 5% of jobs exposed in one of the 90 occupations.

Comparisons based on individual jobs and exposures

To assess whether jobs evaluated with the hybrid method tended to be assigned exposures to more agents, we computed the average and selected quantiles for the number of exposures by job in PROtEuS and in the Lung cancer study. To examine if jobs in PROtEuS were assigned exposures with greater reliability, we compared the relative distribution of the reliability values assigned to the individual exposed job/agent pairs (e.g., each of the 58 combination welder jobs exposed to nitrogen oxides in Figure 2) between the two studies. Since a job could be exposed to several agents, the number of data points in the analysis could be greater than the total number of jobs. We also applied a cumulative logistic model on the reliability ratings of the exposed job/agent pairs, using the study as the predictor with the Lung study representing the reference level. This model provided a quantitative measure of the differences in the relative distribution of reliability levels assigned between the studies as a cumulative odds ratio (OR). A cumulative OR of 2 can be interpreted as a two-fold increase in the odds of an exposed job assigned a reliability level of probable or definite (relative to possible), or assigned a level of definite (relative to possible or probable) in PROtEuS compared to the lung cancer study. This model was also applied to the indices of concentration and frequency to evaluate potential differences in the ratings assigned to exposed jobs between the two studies.

Comparisons based on occupation-agent combinations

We evaluated if the use of the hybrid method tended to increase the proportion of jobs exposed (ie., prevalence) to an agent in an occupation compared to jobs coded with the traditional expert method. To this end, we summarized the exposure data in both studies by combination of occupation and agent. We first assessed if the occupation-agent combinations with non-null exposure in PROtEuS (defined as ≥5% of jobs exposed) were also associated with non-null exposure in the Lung study, representing concordance in exposure status. We then evaluated if the prevalence in PROtEuS was higher compared to the Lung study within the subset of concordant exposed occupation-agent combinations. We also assessed if the trends remained when changing the threshold for non-null exposure to >0%, ≥10%, ≥25% or ≥50%.

To assess if the experts using the hybrid approach tended to assign more frequently the same reliability, intensity and frequency category to jobs exposed to an agent within the same occupation (akin to the application of a JEM), we applied the categorical scheme used to illustrate the variability in exposure in JEPs using colours to the PROtEuS data. The range of relative percentages corresponding to the green colour, representing >75% to 100% of jobs assigned to one categorical rating, was split into one category for >75% to <100% and another for 100% to represent complete homogeneity. We then compared the relative distribution of occupation-agent combinations by category of relative percentage of jobs assigned to one rating in PROtEuS and in the Lung study in the subset of concordant-exposed combinations. Since the proportions for combinations with few jobs can only take a narrow range of values (e.g., 50% or 100% for 2 exposed jobs), the comparison was restricted to occupation-agent combinations with at least 5 exposed jobs in each dataset.

Subanalyses

We conducted additional analyses to evaluate if the trends observed in the distribution of probability and of the categorical ratings of jobs remained after changing the following parameters: (1) Restricted to occupations with a JEP that had undergone expert review to add detailed comments; (2) Stratified by chemical/physical group; (3) Stratified by 2-digit CCDO major group; (4) Stratified by employment period, either over the period with at least 500 blue collar jobs in both studies (1953-1993), or between two periods split at the midpoint of the period covered by the job histories (1934-1972 and 1973-2012).

6.4 Results

Description of exposure assessments based on the hybrid approach

The experts evaluated exposures for a total of 16,065 jobs held by 4005 subjects. Subjects were on average 65 years old at interview (interquartile interval 61-70 years). The number of jobs evaluated by subject ranged from 1 to 13 (average 4), with an average period of 39 years between the first and last job held (interquartile interval 34-44 years). Overall, the review and assignment of exposure for each job description was estimated to take on average under 1 hour. The job histories covered 2263 7-digit occupations, of which 1122 (49.6%) had a JEP available. The remaining 1141 occupations (50.4%) not covered by a specific JEP represented a smaller fraction of all jobs (n=3047, 19%).

A total of 313 agents had at least one job with exposure, Among the 16,065 jobs evaluated, 12,162 (76% of total) were exposed to at least one agent. This proportion was higher among blue collar jobs compared to white collar (89% vs. 58%), as was the average number of agents with exposure by job (10.9 vs. 4.2). Supplementary table II presents the proportion of all jobs exposed to each of the 313 agents (at any reliability level), and stratified by blue/white collar status. Volatile organic liquids had the largest proportion of jobs exposed with 38.7%, followed by alkanes (C5-C17, 26.7%), organic solvents (22.7%) and any polycyclic aromatic hydrocarbons (19.1%). For most agents, the proportion of exposed jobs was higher in blue collar occupations compared to white collar occupations. The few agents (n=28) with a higher prevalence in white collar jobs included inks (+1.8%), calcium sulfate (+2.3%) and calcium carbonate (+5.7%), the latter two mainly associated with teaching occupations from the use of chalk. The experts assigned the reliability of exposure as definite 59% of the time, compared to

28% for probable and 13% for possible; the proportion of definite exposures was slightly higher for blue collar jobs (60%) compared to white collar (55%).

Comparison of exposures assigned in PROtEuS and in the Lung cancer source study

Comparison based on jobs

By comparing the exposure data of jobs in blue collar occupations evaluated in PROtEuS to the Lung cancer study serving as the reference, one endpoint evaluated was if the use of the hybrid method translated into more agents with exposures assigned to jobs. A small increase was found in the average number of exposures by job (among the 203 agents included in the comparison) with 7.7 in PROtEuS jobs (median 5, interquartile interval 2-11) compared to 7.4 for Lung study jobs (median 5, interquartile interval 3-9). Figures were slightly higher in the 75 occupations featuring an expert-annotated JEP with an average of 9.0 agents/job in PROtEuS compared to 8.5 for the reference study. When stratifying the jobs in two periods, the difference was greater for 1934-1972 (average PROtEuS 8.0, Lung 7.2) compared to 1973-2012 (PROtEuS 7.6, Lung 7.7).

The proportion of all jobs (n=4318 for PROtEuS, n=3022 for Lung) exposed to each of the 203 agents retained is presented in Supplementary table II. The rankings of the agents by their prevalence of exposure between the two studies were highly correlated with a Kendall correlation coefficient of 0.81. Diesel engine emissions was the most prevalent agent in PROtEuS and ranked second in Lung, both with 31%. Leaded engine emissions had the highest prevalence among Lung cancer study jobs with 42%, and ranked third for PROtEuS jobs with 29%.

To evaluate whether the experts tended to assign exposure with higher reliability in PROtEuS, we compared the relative distribution of the reliability categories of the exposed job/agent pairs (n=29,551) with those from the Lung cancer study (n=18,864), shown in Figure 3. There was a higher proportion of exposures with definite reliability among PROtEuS jobs (61.4%) compared to the Lung cancer study (55.4%), with fewer at possible (11.7% vs. 15.9%) and probable (27.0% vs. 28.7%) reliability. This trend was reflected in the cumulative OR estimated of 1.31 (95% CI 1.26-1.36). Analyses stratified by agent group, CCDO major group and employment period, presented in Supplementary table III, showed comparable trends except for Sales Occupations, which was associated with lower reliability in PROtEuS (OR 0.81, 95%CI 0.68-0.96). The opposite trend was found for the other metrics with fewer exposures assigned to the higher concentration and frequency categories in PROtEuS compared to Lung study jobs.

Comparison based on occupation-agent combinations

A high level of agreement was found in the prevalence of exposure across the 18,270 occupation-agent combinations: 85.5% of combinations had a prevalence <5% in both PROtEuS and Lung (concordant exposed pairs), and another 8.2% had a prevalence ≥5% in both studies (Table II). Discordance was twice likely to reflect some exposure in the Lung cancer study and none in PROtEuS (4.1% vs. 2.2%), although the percentages were closer when using other thresholds for the minimum prevalence (ranging from >0% to ≥50%), to define the exposure status. There was a general trend of higher prevalence of exposure in PROtEuS among concordant exposed pairs (n=1502), with a median increase of 1.3% relative to the Lung cancer study (interquartile interval -7.6% - +12.5%). The trend of higher prevalence in PROtEuS among concordant exposed pairs was less sensitive to the choice of threshold, except for P>0%

(median difference of -0.17%). Results for stratified analyses by chemical/physical group, CCDO major group or employment period (Supplementary table IV) were generally comparable with some exceptions, such as Gases where the median difference in prevalence was lower in PROtEuS by 1.2%.

The comparisons based on occupation-agent combinations also aimed at evaluating the influence of the exposure assessment method on the variability in the categorical ratings of exposed jobs within an occupation-agent combination. As shown in Figure 4, there was a clear trend of more jobs assigned to the same rating in PROtEuS compared to exposures coded using the traditional method for all metrics. For example, 8.4% of combinations had all jobs assigned the same reliability rating in Lung, compared to 17.1% in PROtEuS. This effect was larger for concentration where all jobs were assigned the same category in 31.1% of combinations for the Lung cancer study and 59.5% for PROtEuS.

6.5 Discussion

The hybrid expert approach applied descriptive summaries of existing exposure evaluations as a tool to shorten the expert-based retrospective exposure assessment for a large number of agents in a community-based study, without compromising on quality. It benefits from the positive attributes of a JEM by streamlining homogeneous exposure profiles while maintaining the main asset of the expert method by accounting for the idiosyncrasies of specific jobs.

The feedback of the PROtEuS experts using this method has been positive, especially from the more senior experts who had a long history of applying the traditional approach, as the hybrid method provided them with more structure, guidance, and readily accessible source of information, yet still allowing them complete latitude in their assignments. The JEPs also represented a useful tool for training junior experts. The impact of this approach was also reflected in an estimated two-fold reduction in the experts' time to carry out their assignment of job/industry titles and exposures, from an average of 2 hours/job in previous studies, to less than one hour in PROtEuS. One source of time saving relates to the color schemes used in the JEPs, whereas green colorings were associated with highly homogeneous assignments in earlier studies that allowed experts to invest more efforts towards more heterogeneous exposures between jobs. The larger proportion of jobs in white collar occupations in PROtEuS compared to earlier studies is another factor that likely shortened the overall coding time as they tended to be associated with fewer exposures.

The use of past exposure evaluations was also applied in the Lung cancer study, albeit informally, where the experts had access to crude summaries of the exposures assigned in the previous Multisite study by occupation (with less precise 4-digit CCDO codes). They also occasionally used a small sample of jobs within an occupation to inform their judgment. However, their use of these exposure sources tended to be inconsistent for lack of a comprehensive computerized interface. In contrast, the hybrid approach represents a systematic and coordinated application of available as well as enhanced exposure information in a study.

Comparison of exposures assigned in PROtEuS and in a source study

The comparison performed to evaluate the exposures assigned with the hybrid method to those coded with the traditional expert approach was not designed to appraise its reliability, where a comparison based on independent assessment of the same jobs using both approaches would have been more suited, such as in previous evaluations of the Montreal expert method (Goldberg *et al.*, 1986; Siemiatycki *et al.*, 1997). Another approach would be to assess the performance of the hybrid method when applied to a job description for which industrial measurements are available (Fritschi *et al.*, 2003). Our group has also applied a simplified version of the hybrid approach to evaluate occupational exposure to engine emissions in other population studies in Canada, where expected exposure-cancer associations were observed (Latifovic *et al.*, 2015; Kachuri *et al.*, 2016).

In the comparison between the exposures assigned to a subset of blue collar jobs with either approaches, the average number of agents with exposure by job was slightly higher in PROtEuS, although the difference was small (difference of 0.3). The trend was clearer for reliability where the experts tended to more frequently assign exposures as definite with the hybrid method, owing largely to the wealth of information available in JEPs. In the case of concentration and frequency of exposure, jobs in PROtEuS were less likely to be assigned to lower categories (eg., low concentration). The differences between studies for these two metrics tended to be more variable between the chemical groups, compared to reliability where the overall trend generally remained consistent in stratified analyses,

The high concordance in exposure status in analyses stratified by occupation-agent combination suggests that when there was some exposure in JEP, experts were likely to assign exposure in PROtEuS, and the opposite for no exposure. For the discordance in exposure status, no clear trend emerged since the difference between the proportion of combinations with some exposure only in the Lung cancer study or in PROtEuS varied depending on the threshold used on minimum prevalence to define exposure.

The hypothesis of lower heterogeneity in the exposures assigned to jobs within an occupation from the use of the hybrid method was confirmed when analyzing the categorical exposure indices of exposure. This trend was stronger for concentration where nearly 60% of occupation-agent combinations had all jobs exposed at the same level in PROtEuS (31% in the Lung cancer study). The trend towards more jobs being assigned the same rating can be interpreted in two contrasting ways. On one hand, this may represent higher coherence in the exposures assigned by the experts, which may partly result from the comments in the JEPs. On the other hand, this may also result from the experts putting a higher weight on past data in their judgment compared to the specificities of the individual job descriptions. Despite the patterns observed, there generally remained significant variability in the ratings assigned for most occupation-agent combinations, suggesting that the experts integrated both sources of information in their assessments.

While the comparisons aimed to evaluate differences in the exposures assigned using two related methods, they were performed on jobs from two different study populations, distributed in different occupations, and held at somewhat different times which may confound the trends observed. However, both studies were conducted in the same region, and some of these differences were mediated by restricting the comparisons to a set of common agents and occupations. Moreover, the stratification by employment period in sensitivity analyses yielded comparable results.

Limitations

The comparisons of exposures assigned to jobs using the two approaches covered less than 10% of all blue collar 7-digit CCDO occupational categories in the job histories of the PROtEuS and Lung cancer studies. The 90 occupations included were however highly prevalent, representing approximately half of all blue-collar jobs evaluated. In addition, some of the trends observed may partly result from differences in the distribution of jobs across occupations between the two studies. While we restricted the comparisons to blue collar occupations with the most data in common between the two studies for this reason, some residual differences in the distribution of jobs by occupation in this subset may remain.

The development of the hybrid method was contingent of the availability of a large pool of exposure data spanning a wide range of occupations in the population, and of specialized expertise to interpret and augment this information with comments and guidelines. In our case, we could source data from two large case-control studies conducted in the same region, and two experts had over 20 years of experience in implementing the traditional expert method. However, significant data gaps remained since half of the occupations encompassed by PROtEuS jobs did not have a specific JEP available. These represented less prevalent occupations, and the additional exposure data collected for these jobs could improve the coverage of occupations in future studies using the hybrid method. Regarding the difficulty in finding data to create profiles, the recent CANJEM matrix (available from www.canjem.ca) summarizing expert evaluations from studies conducted in Montreal (including the two source studies of JEPs) may constitute an initial source of information to implement the hybrid method for other investigators.

6.6 Conclusion

The application of the hybrid expert method decreased the time needed to evaluate exposures and resulted in a small increase in the number and level of reliability of exposures assigned to jobs compared to jobs coded with the traditional expert method. It also reduced the variability in the ratings assigned to jobs exposed to an agent within the same occupation, although whether this reflects greater coding coherence or over-influence of JEPs is unclear.

The hybrid method is one of a few recent strategies to increase the efficiency of retrospective exposure assessment in population studies such as predefined decision rules (Fritschi *et al.*, 2009) or evaluations restricted to a sub-sample of all jobs, or "two-phase" method (Wild *et al.*, 2016), that slots between the assessment based solely on individual job descriptions, and the group-based assignment of JEMs. Among those, the hybrid method leans closer to the traditional expert method as it involves the review of each job description, although guided by past data. Finally, a valuable advantage of the hybrid method is the greater transparency in the assessment represented by the exposure information assembled and the overall coding rules used by the experts to assign exposures.

Acknowledgments

The coding of exposures was carried out by Michel Gérin, Louise Nadon, Denis Bégin, Ramzan Lakhani, Benoît Latreille, Mounia Rhazi and Robert Bourbonnais. Louise Nadon and Ramzan Lakhani were also instrumental in the development of the hybrid method and the organization of the coding infrastructure in PROtEuS. The contribution of Hugues Richard in the preparation of the exposure data is gratefully acknowledged.

6.7 Tables and figures

Table I. Selected characteristics of the PROtEuS study and the earlier Lung and Multisite cancer studies

	PROtEuS	Lung cancer study (male subjects)	Multisite cancer study
Number of subjects	4013	1661	4263
Cases	*1966 (49.0%)*	*762 (45.9%)[1]*	*3730[2]*
Controls	*2047 (51.0%)*	*899 (54.1%)*	*533*
Years conducted	2005-2009	1996-2001	1979-1986
Number of jobs	16,065	6881	15,067
Blue collar[3]	*9239 (57.5%)*	*5381 (78.2%)*	*11,468 (76.1%)*
White collar[4]	*6826 (42.5%)*	*1500 (21.8%)*	*3597 (23.9%)*
Period covered by jobs	1943-2012	1934-1999	1920-1986

1. Includes 24 mesothelioma cases
2. Cases covered 19 different cancer sites
3. Jobs in 4-digit CCDO unit groups classified as skilled, semiskilled, unskilled and farming occupations in the classification of Pineo *et al.* (1977)
4. Jobs in 4-digit CCDO unit groups classified as professionals, management, technicians, supervisors and foremen in the classification of Pineo *et al.* (1977)

Table II. Agreement in exposure status among occupation-agent combinations (n=18,270) between Lung and PROtEuS jobs based on different thresholds for prevalence defining exposure.

	Percent concordant		Percent discordant		Concordant exposed occupation-agent combinations		
Prevalence threshold (1)	Exposed (%)	Unexposed (%)	Exposed Lung (%) (2)	Exposed PROtEuS (%) (3)	Number of combinations (4)	Kendall correlation in probability (5)	Median difference. in probability, PROtEuS-Lung (6)
P>0%	11.8	76.8	6.0	5.4	2162	0.61	-0.17
P≥5%	8.2	85.5	4.1	2.2	1502	0.55	1.34
P≥10%	6.2	90.2	2.2	1.4	1127	0.52	2.35
P≥25%	3.7	94.2	0.9	1.1	685	0.43	2.90
P≥50%	2.2	96.3	0.6	0.9	403	0.36	2.27

1. Minimum prevalence (proportion of jobs exposed) in an occupation-agent combination to be considered exposed
2. Proportion of occupation-agent combinations with probability ≥5% in Lung and <5% in PROtEuS
3. Proportion of occupation-agent combinations with probability ≥5% in PROtEuS and <5% in Lung
4. Number of concordant exposed occupation-agent combinations (probability ≥5% in Lung and ≥5% in PROtEuS)
5. Kendall correlation in the probability of exposure between concordant exposed occupation-agent combinations
6. Median difference in probability (probability in PROtEuS minus probability in Lung) across concordant exposed occupation-agent combinations

Figure 1. Development of the job-exposure profiles from the source studies and their application in PROtEuS, and selection of the exposure data from jobs in blue-collar occupations used for the comparisons

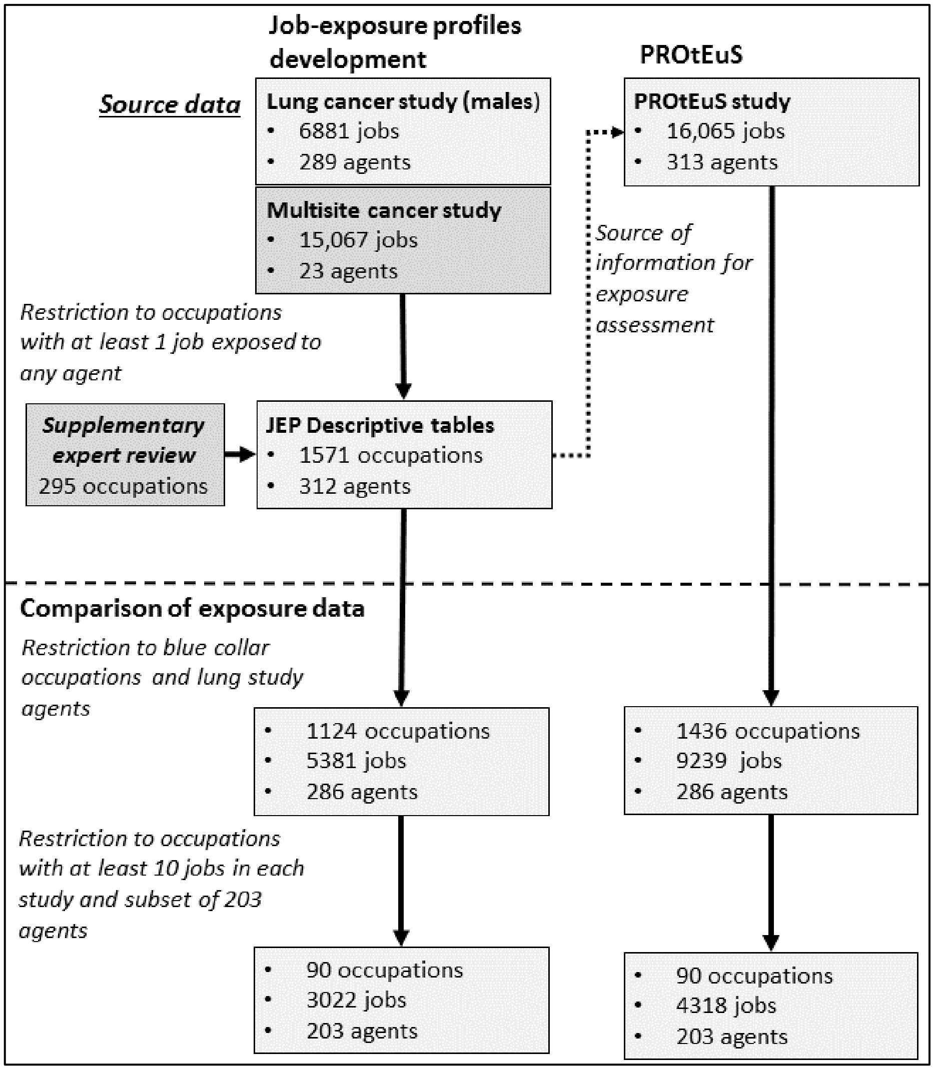


Figure 2. Subset of the JEP for Combination welders (CCDO 8335-126) with 5 out of 111 agents, based on 37 jobs from the Lung cancer study and 61 jobs from the Multisite cancer study.

Agent	N	Reliability			Concentration			Frequency			Comments
		Possible	Probable	Definite	Low	Medium	High	<5%	5-30%	>30%	
*Iron fumes	58		1	**57**	2	20	**36**	1	10	**47**	This is the most likely base metal in construction steel AND stainless steel. High concentration in most cases
*Manganese fumes	57		1	**56**	21	**36**		2	10	**45**	Almost all commercial steel contains manganese which is introduce to deoxidize, desulfurize and to add strength. Medium concentration, low if environmental
Nitrogen oxides	37			**37**	**31**	6			1	**36**	The ultraviolet light of the arc produces nitrogen oxides (NO, NO2), from the nitrogen (N) and oxygen (O2) in the air.
Ozone	37		10	**27**	**34**	3			11	**26**	Ozone (O3) is produced by ultraviolet light from the welding arc.
Mild steel dust	31	2	8	**21**	**14**	11	6	2	**15**	14	Sanding for surface preparation and grinding to smooth the weld

Legend
>75% of jobs with rating
≥50-75% of jobs with rating
<50% of jobs with rating

* Denotes agent present on the checklist of the Multisite cancer study. N: Number of jobs exposed to the agent

Figure 3. Relative distributions of the exposures assigned to jobs by categorical exposure metric in the Lung cancer study and PROtEuS

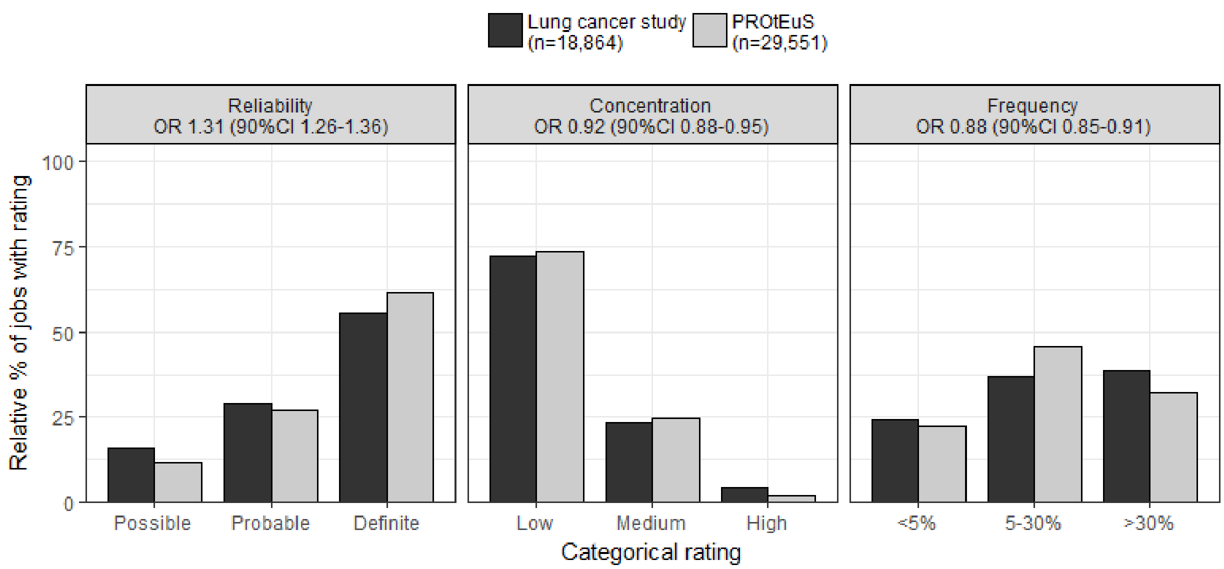
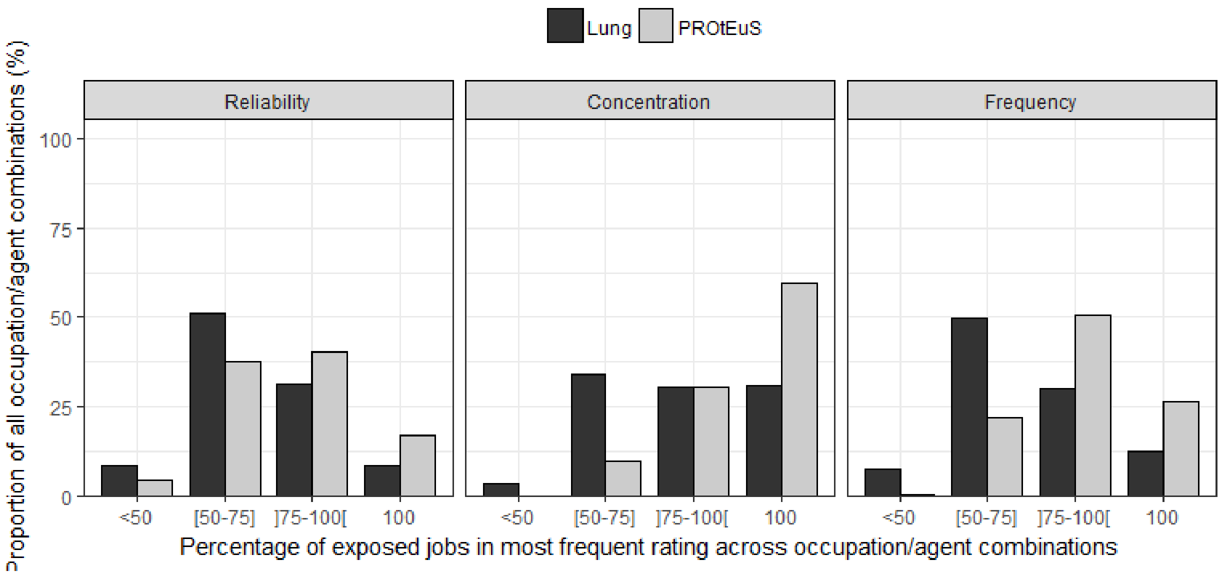


Figure 4. Distribution of the proportion of jobs in the most frequently assigned rating across concordant exposed occupation-agent pairs with at least 5 exposed jobs (n= 926) between Lung and PROtEuS, stratified by exposure metric



6.8 References

Aronson KJ, Howe GR, Carpenter M *et al.* (2000) Occupational Surveillance in Canada: Cause-Specific Mortality Among Workers, 1965-91. (Statstics Canada Catalogue No. 84-546-XCB). Minister of Industry. Ottawa: Statistics Canada.

Behrens T, Mester B, Fritschi L. (2012) Sharing the knowledge gained from occupational cohort studies: a call for action. Occup Environ Med; 69: 444-48.

Blanc-Lapierre A, Spence A, Karakiewicz PI *et al.* (2015) Metabolic syndrome and prostate cancer risk in a population-based case–control study in Montreal, Canada. BMC Public Health; 15: 913.

Bouyer J, Hémon D. (1993) Retrospective Evaluation of Occupational Exposures in Population-Based Case-Control Studies: General Overview with Special Attention to Job Exposure Matrices. Int J Epidemiol; 22: S57-S64.

Department of Employment and Immigration. (1971) Canadian classification dictionnary of occupations, vol. 1. Ottawa, ON: Department of Employment and Immigration.

Friesen MC, Wheeler DC, Vermeulen R *et al.* (2016) Combining Decision Rules from Classification Tree Models and Expert Assessment to Estimate Occupational Exposure to Diesel Exhaust for a Case-Control Study. Ann Occup Hyg; 60: 467-78.

Fritschi L, Friesen MC, Glass D *et al.* (2009) OccIDEAS: Retrospective Occupational Exposure Assessment in Community-Based Studies Made Easier. J Environ Public Health; 2009, Article ID: 957023.

Fritschi L, Nadon L, Benke G *et al.* (2003) Validation of expert assessment of occupational exposures. Am J Ind Med; 43: 519-22.

Fritschi L, Sadkowsky T, Benke GP *et al.* (2012) Triaging Jobs in a Community-Based Case–Control Study to Increase Efficiency of the Expert Occupational Assessment Method. Ann Occup Hyg; 56: 458-65.

Gérin M, Siemiatycki J, Kemper H *et al.* (1985) Obtaining occupational exposure histories in epidemiologic case-control studies. J Occup Med; 27: 420-6.

Goldberg MS, Siemiatycki J, Gérin M. (1986) Inter-rater agreement in assessing occupational exposure in a case-control study. Br J Ind Med; 43: 667-76.

Kachuri L, Villeneuve PJ, Parent M-É *et al.* (2016) Workplace exposure to diesel and gasoline engine exhausts and the risk of colorectal cancer in Canadian men. Environ Health; 15: 4.

Kauppinen T, Toikkanen J, Pukkala E. (1998) From cross-tabulations to multipurpose exposure information systems: A new job-exposure matrix. Am J Ind Med; 33: 409-17.

Kromhout H, Vermeulen R. (2001) Application of job-exposure matrices in studies of the general population-some clues to their performance. European Respiratory Review; 11: 80-90.

Latifovic L, Villeneuve PJ, Parent ME *et al.* (2015) Bladder cancer and occupational exposure to diesel and gasoline engine emissions among Canadian men. Cancer Med; 4: 1948-62.

Parent M-É, Rousseau M-C, Boffetta P *et al.* (2007) Exposure to Diesel and Gasoline Engine Emissions and the Risk of Lung Cancer. Am J Epidemiol; 165: 53-62.

Pineo PC, Porter J, McRoberts HA. (1977) The 1971 Census and the Socioeconomic Classification of Occupations. Canad Rev Sociol Anthrop; 14: 91-102.

Pintos J, Parent M-É, Richardson L *et al.* (2012) Occupational exposure to diesel engine emissions and risk of lung cancer: evidence from two case–control studies in Montreal, Canada. Occup Environ Med; 69: 787-92.

Pronk A, Stewart PA, Coble JB *et al.* (2012) Comparison of two expert-based assessments of diesel exhaust exposure in a case–control study: programmable decision rules versus expert review of individual jobs. Occup Environ Med; 69: 752-58.

Siemiatycki J. (1991) Risk factors for cancer in the workplace. Boca Raton: CRC Press.

Siemiatycki J, Fritschi L, Nadon L *et al.* (1997) Reliability of an expert rating procedure for retrospective assessment of occupational exposures in community-based case-control studies. Am J Ind Med; 31: 280-86.

Statistics Canada. (1980) Standard Industrial Classification, 1980. Ottawa: Statistics Canada, Standards Division.

Teschke K. (2003) Exposure surrogates: job-exposure matrices, self-reports, and expert evaluations. In Nieuwenhuijsen, MJ. Exposure Assessment in Occupational and Environmental Epidemiology, Oxford: Oxford University Press.

Teschke K, Olshan AF, Daniels JL *et al.* (2002) Occupational exposure assessment in case-control studies: opportunities for improvement. Occup Environ Med; 59: 575-94.

Vida S, Pintos J, Parent ME *et al.* (2010) Occupational exposure to silica and lung cancer: pooled analysis of two case-control studies in Montreal, Canada. Cancer Epidemiol Biomarkers Prev; 19: 1602-11.

Wheeler DC, Burstyn I, Vermeulen R *et al.* (2013) Inside the black box: starting to uncover the underlying decision rules used in a one-by-one expert assessment of occupational exposure in case-control studies. Occup Environ Med; 70: 203-10.

Wild P, Schil lW, Bourgkard E *et al.* (2016) The 2-phase case–control design: an efficient way to use expert-time. Scand J Work Environ Health; 42: 162-69.

6.9 Appendices

Supplementary table I. List of the 90 blue collar occupations included in the analysis

	Number of jobs	
CCDO code and title	**Lung**	**PROtEuS**
3135114: ORDERLY (medical)*	13	56
4131134: ACCOUNTING CLERK (clerical)	54	43
4135182: UTILITY CLERK, BANK (bank. & finance)	14	10
4153118: SHIPPING AND RECEIVING CLERK (clerical)	39	43
4153126: SHIPPING CLERK (clerical)	37	76
4155126: STOREKEEPER (clerical)	23	89
4155146: INVENTORY CLERK (clerical)	12	23
4172110: LETTER CARRIER (gov. serv.)*	26	32
4177118: MESSENGER (clerical)	14	15
4177122: DELIVERY PERSON (ret. trade)	74	104
4197114: CLERK, GENERAL OFFICE (clerical)	64	78
5133126: SALES REPRESENTATIVE, COMMERCIAL AND INDUSTRIAL EQUIPMENT AND SUPPLIES (whole. trade)*	15	51
5133130: SALES REPRESENTATIVE, FOOD PRODUCTS (whole. trade)*	26	22
5133199: OTHER COMMERCIAL TRAVELLERS*	15	72
5135110: SALESPERSON, MOTOR VEHICLES (ret. trade)*	21	28
5135154: SALESPERSON, HARDWARE (ret. trade; whole. trade)*	13	52
5137111: SUPERMARKET CLERK (ret. trade)	22	54
5137114: SALES CLERK (ret. trade)	74	99
5145110: SERVICE-STATION ATTENDANT (motor vehicle; ret. trade)*	23	26
5172118: SALESPERSON, REAL ESTATE (insur. & real estate)	26	73
5193118: ROUTE DRIVER (any ind.)*	70	46
6112158: POLICE OFFICER (gov. serv.)*	19	95
6115138: SECURITY GUARD (any ind.)	85	106
6117190: INFANTRY SOLDIER (military)*	75	74
6121127: COOK, FIRST (cater. & lodg.)*	25	10
6121130: SHORT-ORDER COOK (cater. & lodg.)*	13	14
6121134: COOK, THIRD (cater. & lodg.)*	16	24
6123110: BARTENDER (cater. & lodg.)*	23	28
6125126: WAITER/WAITRESS (cater. & lodg.)*	71	72
6165126: PRESSER, MACHINE (garment & fabric; laund., clean. & press.)*	28	52
6191110: JANITOR (any ind.)*	87	140
6191114: CLEANER, light DUTY (any ind.)*	36	50
6191118: CLEANER, INDUSTRIAL-PLANT (any ind.)	18	14
6198134: KITCHEN HELPER (cater. & lodg.)*	10	16
6198170: DISHWASHER (cater. & lodg.)*	15	32

CCDO code and title	Number of jobs	
	Lung	PROtEuS
7111110: FARMER, GENERAL (agric.)*	22	37
7181110: FARM WORKER, GENERAL (agric.)*	94	26
7183122: FARM WORKER, VEGETABLE (agric.)	10	12
7195146: LANDSCAPE WORKER (agric.)*	14	15
7198112: FARM LABOURER, GENERAL (agric.)*	29	28
7513122: LOGGER, ALL-ROUND (forest. & log.)*	88	58
8213114: BAKER (bake. prod.; cater. & lodg.; food & bev., n.e.c.)*	20	57
8215110: BUTCHER, ALL-ROUND (slaught. & meat pack.)*	26	52
8215114: MEAT CUTTER (ret. trade; slaught. & meat pack.)*	10	18
8238134: SAWMILL LABOURER (pulp & paper; sawmill; woodworking)*	12	13
8311110: TOOL AND DIE MAKER (mach., weld. & forg.)*	14	13
8313154: MACHINIST, GENERAL (mach., weld. & forg.)*	36	106
8333118: SHEET-METAL WORKER (construction; metal stamp., press. & coat.)*	19	35
8335126: WELDER, COMBINATION (mach., weld. & forg.)*	37	53
8335138: WELDER, ARC (mach., weld. & forg.)*	22	59
8541110: CABINETMAKER (furn.)*	24	60
8551146: CUTTER, PORTABLE MACHINE (garment & fabric)*	18	29
8553110: TAILOR, MADE-TO-MEASURE GARMENTS (garment & fabric; ret. trade)*	18	28
8563114: SEWING-MACHINE OPERATOR (any ind.)*	28	15
8581110: MOTOR-VEHICLE MECHANIC (motor vehicle)*	73	82
8581118: INDUSTRIAL-TRUCK MECHANIC (mech. equip., n.e.c.)*	17	19
8581142: BODY REPAIRER (motor vehicle)*	21	61
8582110: AIRCRAFT MECHANIC (air & space-craft)*	15	22
8584122: MILLWRIGHT (mech. equip., n.e.c.)*	14	74
8711110: HEAVY-DUTY-EQUIPMENT OPERATOR (any ind.)*	36	29
8733122: ELECTRICIAN (construction)*	41	111
8781110: CARPENTER (construction)*	65	104
8781121: CONCRETE FORMER (construction)*	10	18
8782110: BRICKLAYER (construction)*	15	18
8785110: PAINTER AND DECORATOR (construction)*	14	24
8785120: PAINTER (construction)*	30	27
8787118: ROOFER, ASPHALT (construction)*	10	10
8791110: PIPE FITTER (construction)*	14	27
8791114: PLUMBER (construction)*	14	56
8793114: STRUCTURAL-STEEL ERECTOR (construction)*	13	22
8798114: CONSTRUCTION LABOURER (construction)*	148	73
9171110: BUS DRIVER (motor trans.)*	34	102
9173110: TAXI DRIVER (motor trans.)*	70	105
9173114: CHAUFFEUR (motor trans.)*	13	31

CCDO code and title	Number of jobs	
	Lung	PROtEuS
9175110: TRUCK DRIVER, GENERAL (motor trans.)*	130	176
9175114: DRIVER, TANK TRUCK (motor trans.)*	19	26
9175118: TRUCK DRIVER, HEAVY (motor trans.)*	65	87
9175122: TRUCK DRIVER, TRACTOR-TRAILER (motor trans.)*	12	15
9175130: DRIVER, DUMP-TRUCK (motor trans.)*	22	16
9175138: TRUCK DRIVER, LIGHT (motor trans.)*	62	54
9179190: TRUCK-DRIVER HELPER (motor trans.)*	14	22
9313110: LONGSHORE WORKER (water trans.)*	24	20
9315126: INDUSTRIAL-TRUCK OPERATOR (any ind.)*	33	101
9317218: PACKAGER, MACHINE (any ind.)*	14	17
9318110: MATERIAL HANDLER, GENERAL (any ind.)*	68	73
9318114: MATERIAL HANDLER, HEAVY (any ind.)*	33	22
9318118: MATERIAL HANDLER, LIGHT (any ind.)*	33	28
9318122: PACKAGER, HAND (any ind.)*	35	51
9512126: OFFSET PRESS OPERATOR (print. & pub.)*	15	37
9918110: LABOURER, MUNICIPAL (gov. serv.)*	39	25

* Denotes an occupation with an expert-annotated job-exposure profile. n.e.c.: Not elsewhere classified

Supplementary table II. Proportion of jobs exposed by agent in PROtEuS (n=313), overall and stratified by blue/white collar status, and proportion of PROtEuS and Lung cancer study jobs exposed among the blue collar occupations (n=90) and agents (n=209) retained in the comparison

	Global prevalence in PROtEuS (%)			Prevalence in blue-collar jobs in comparison (%)	
Group/agent	All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	PROtEuS[4]	Lung[5]
Metals					
Metallic dust	13.8	18.3	7.6	17.5	12.6
Bronze dust	0.4	0.6	0.1	0.9	0.9
Brass dust	0.4	0.6	0.2	0.6	0.7
Stainless steel dust	1.9	2.7	0.8	3.3	3.5
Mild steel dust	8.9	12.7	3.8	13.0	10.6
Inorganic pigments	5.1	6.9	2.7	6.4	7.2
Alumina	7.0	10.7	2.0	13.2	8.6
Titanium dioxide	0.8	1.3	0.2	2.3	3.0
Iron oxides	3.1	4.8	0.8	6.0	5.8
Zinc oxide	0.2	0.3	0.0	0.3	0.2
Lead oxides	0.5	0.7	0.1	0.9	0.3
Basic lead carbonate	0.3	0.5	0.0	0.9	0.4
Lead chromate	1.1	1.8	0.3	2.8	1.3
Gas welding fumes	5.2	7.2	2.5	8.2	6.9
Arc welding fumes	5.4	7.1	3.2	8.5	7.4
Soldering fumes	4.3	5.7	2.4	6.0	3.3
Metal oxide fumes	9.3	12.4	5.1	12.9	9.9
Chromium (VI)	2.7	4.0	0.9		
Beryllium	0.1	0.0	0.1		
Magnesium	0.1	0.2	0.0		
Aluminum	8.5	12.6	3.0		
Titanium	1.2	1.9	0.3		
Vanadium	0.0	0.0	0.0		
Chromium	4.4	6.2	1.9		
Manganese	4.4	6.4	1.6		
Iron	11.6	16.3	5.2		
Cobalt	0.2	0.2	0.1	0.2	0.1
Nickel	2.5	3.5	1.2		
Copper	3.9	5.9	1.1		

Group/agent	Global prevalence in PROtEuS (%) All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	Prevalence in blue-collar jobs in comparison (%) PROtEuS[4]	Lung[5]
Zinc	2.9	4.3	1.0		
Arsenic	0.5	0.8	0.1	0.3	0.8
Selenium	0.6	0.9	0.1	1.5	1.9
Silver	0.8	0.9	0.7		
Cadmium	1.9	2.7	0.8		
Tin	3.5	4.6	1.9		
Antimony	0.2	0.4	0.0	0.0	0.1
Tungsten compounds	0.2	0.4	0.0		
Gold compounds	0.0	0.0	0.0		
Mercury	0.4	0.3	0.6		
Lead	8.4	12.5	2.9		
Organic Solvents					
n-Hexane	1.4	1.9	0.7	2.8	1.7
Methanol	2.4	2.0	3.0	2.7	4.8
Ethanol	0.6	0.6	0.7	0.2	0.5
Ethylene glycol	2.3	3.1	1.1	4.1	3.8
Isopropanol	4.0	5.3	2.2	7.6	7.4
Diethyl ether	0.2	0.1	0.3	0.1	0.1
Carbon tetrachloride	0.6	0.9	0.2	0.2	0.2
Chloroform	0.2	0.1	0.3		
Methylene chloride	0.4	0.6	0.2	0.4	0.8
1,1,1-Trichlorethane	0.4	0.6	0.2	0.7	0.5
Trichloroethylene	0.5	0.6	0.3	0.7	0.5
Perchloroethylene	0.2	0.4	0.1	0.3	0.5
Acetone	1.5	1.3	1.7	0.9	0.6
Methyl ethyl ketone	0.0	0.0	0.0		
Benzene	4.7	6.8	1.8		
Toluene	4.9	6.8	2.2		
Xylene	4.0	5.7	1.6		
Styrene	1.0	1.6	0.1	1.5	0.7
Organic Solvents	22.7	28.3	15.3	27.4	25.9
Mineral spirits post 1970	8.3	11.7	3.8	13.2	7.0
Mineral spirits pre 1970	5.4	8.5	1.3	8.6	8.1
Aliphatic alcohols	7.9	8.8	6.7		
Chlorinated alkanes	2.0	2.7	1.1		
Chlorinated alkenes	1.6	2.2	0.9		

Group/agent	Global prevalence in PROtEuS (%)			Prevalence in blue-collar jobs in comparison (%)	
	All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	PROtEuS[4]	Lung[5]
Aliphatic esters	0.8	1.1	0.5		
Aliphatic ketones	2.4	2.7	2.0		
Glycol ethers	0.2	0.3	0.2	0.3	0.3
Mononuclear aromatic hydrocarbons	14.5	21.1	5.7		
Aromatic alcohols	0.8	1.0	0.6		
Polycyclic aromatic hydrocarbons					
PAHs from any source	19.1	27.6	7.5		
PAHs from other sources	3.5	5.3	1.2		
PAHs from wood	1.5	2.4	0.2		
PAHs from petroleum	16.3	23.8	6.2		
PAHs from coal	1.5	2.1	0.6		
Benzo[a]pyrene	4.9	7.5	1.4		
Paints, varnishes and inks					
Other paints,varnishes	0.0	0.0	0.0		
Wood varnishes, stains	0.0	0.0	0.0		
Inks	3.2	2.4	4.2	1.2	1.3
Metal coatings	2.2	3.1	1.0	3.1	3.4
Wood stains	0.7	1.1	0.2	1.5	1.4
Wood varnishes	1.1	1.5	0.5	1.9	1.8
Wood paints	1.1	1.6	0.5	2.0	2.4
Gypsum / Plaster coatings	1.8	2.3	1.2	3.7	4.5
Water based coatings	1.7	1.9	1.4	3.0	3.3
Solvent based coatings	4.4	6.1	2.2	6.8	7.4
Artistic paints	0.1	0.0	0.3		
Combustion products and engine emissions					
Ashes	1.6	2.5	0.4	2.8	4.6
Other pyrolysis fumes	0.0	0.0	0.0		
Cooking fumes	3.5	5.1	1.3	5.0	5.2
Leaded engine emissions	16.1	21.3	9.1	29.3	42.3
Coal combustion products	0.6	0.9	0.2	0.6	1.2
Diesel engine emissions	17.7	23.3	10.2	31.4	31.0
Liquid fuel combustion products	1.6	2.5	0.5	2.1	2.4
Wood combustion products	1.4	2.4	0.2	2.3	3.4
Natural gas combustion products	2.9	3.9	1.5	3.4	2.5
Jet fuel engine emissions	0.8	0.5	1.2	0.5	0.3
Propane engine emissions	2.0	2.8	1.0	3.3	5.0

Group/agent	Global prevalence in PROtEuS (%)			Prevalence in blue-collar jobs in comparison (%)	
	All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	PROtEuS[4]	Lung[5]
Plastics pyrolysis fumes	1.2	1.6	0.7	0.5	1.1
Rubber pyrolysis fumes	1.2	1.9	0.3	1.1	0.6
Propane combustion products	3.4	4.7	1.5	5.1	3.6
Coke combustion products	0.0	0.1	0.0	0.0	0.1
Unleaded engine emissions	16.7	17.6	15.5	23.5	14.2
Paint pyrolysis fumes	2.0	3.1	0.5	3.7	4.0
Mineral-based oil & grease pyrolysis fumes	5.2	7.9	1.6	8.2	6.2
Diesel engine emissions (heavy)	0.6	0.7	0.4	0.0	0.4
Jet fuel engine emissions (wide-cut)	0.1	0.0	0.1		
Diesel engine emissions (light)	17.6	23.2	10.1	31.4	30.9
Petroleum products					
Paraffin	0.0	0.0	0.0		
Leaded gasoline	6.6	8.7	3.9	11.6	16.1
Kerosene	0.5	0.8	0.1	1.3	2.5
Diesel oil	3.2	5.2	0.4	7.2	4.0
Heating oil	0.9	1.3	0.3	1.3	1.3
Crude petroleum	0.2	0.2	0.2	0.1	0.2
Asphalt	1.1	1.4	0.7	2.1	2.9
Other mineral oils	0.7	1.0	0.2	1.0	0.7
Jet fuel (JP5, Jet A, Jet A1)	0.4	0.4	0.4	0.3	0.2
Aviation gasoline	0.3	0.3	0.3	0.4	0.3
Unleaded gasoline	8.9	9.2	8.5	12.0	5.6
Textile oils	0.4	0.7	0.0	0.0	0.1
Jet fuel (wide-cut) (JP4, Jet B)	0.0	0.0	0.0		
Bunker C heating oil	0.0	0.1	0.0		
Alkanes (C18+)	16.2	23.6	6.1		
Alkanes (C5-C17)	26.7	34.6	16.1		
Carbon dust					
Coal dust	0.3	0.4	0.1	0.3	0.8
Carbon black	1.6	1.5	1.8	1.0	1.1
Coke dust	0.0	0.1	0.0	0.0	0.2
Graphite dust	0.2	0.3	0.1		
Soot					
Soot from any source	2.5	4.0	0.5		
Petroleum soot	1.9	2.9	0.5		
Wood Soot	0.5	0.8	0.0	0.1	0.2

Group/agent	Global prevalence in PROtEuS (%) All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	Prevalence in blue-collar jobs in comparison (%) PROtEuS[4]	Lung[5]
Coal soot	0.1	0.1	0.0		
Cutting fluids					
Cutting fluids	0.1	0.1	0.1		
Cutting fluids pre-1955 (straight, mineral-based)	0.1	0.2	0.1	0.1	1.0
Cutting fluids post-1955 (straight, mineral-based)	2.1	2.9	1.1	3.8	2.7
Cutting fluids (emulsified, mineral-based)	1.5	2.2	0.5	2.9	1.1
Cutting fluids (synthetic)	0.0	0.1	0.0	0.1	0.0
Synthetic oils and greases					
Silicone oils and greases	0.0	0.1	0.0		
Synthetic oils and greases	0.0	0.0	0.0		
Complex organic fluids					
Animal, vegetable glues	0.2	0.3	0.0	0.1	0.4
Linseed oil	1.1	1.6	0.4	2.1	2.6
Polyvinyl Acetate	0.0	0.0	0.0		
Synthetic adhesives	5.9	8.0	3.0	7.8	5.9
Waxes, polishes	1.7	2.6	0.3	3.1	2.5
Lubricating oils and greases (mineral-based)	10.8	16.3	3.5	16.4	11.9
Coal tar and pitch	0.6	0.7	0.4	1.0	1.3
Hydraulic fluid	3.0	4.5	0.9	4.4	4.2
Ink formulation oils	0.6	0.8	0.4	0.8	0.7
Heat transfer oils	0.0	0.0	0.0		
Lubricating oils and greases (synthetic)	0.1	0.1	0.1		
Gases					
Hydrogen	0.1	0.2	0.0		
Carbon monoxide	11.1	16.3	4.0		
Hydrogen cyanide	0.6	1.0	0.1		
Ammonia	5.0	6.2	3.4		
Nitrogen oxides	8.9	13.2	3.1		
Ozone	5.9	6.5	5.1	7.9	6.3
Sulphur dioxide	2.7	3.5	1.7		
Chlorine	0.8	1.0	0.5	0.1	0.3
Chlorine dioxide	0.2	0.2	0.3	0.0	0.4
Natural gas	0.4	0.4	0.4	0.3	0.3
Methane	1.5	2.0	0.7		

Group/agent	Global prevalence in PROtEuS (%)			Prevalence in blue-collar jobs in comparison (%)	
	All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	PROtEuS[4]	Lung[5]
Propane	1.5	2.0	0.8	2.2	1.9
Formaldehyde	6.3	8.7	2.9	7.7	10.0
Ethylene oxide	0.1	0.0	0.1		
Ethylene	0.0	0.0	0.0		
Acetylene	2.6	3.9	0.9	4.8	3.3
Vinyl chloride	0.3	0.5	0.1	0.1	0.2
Phosgene	0.5	0.9	0.0	0.2	0.1
Anaesthetic gases	0.4	0.3	0.6	0.3	0.0
Propellant gases	1.7	2.6	0.7	1.4	1.0
Coal gas	0.0	0.0	0.0		
Trichlorotrifluoroethane or CFC-113	0.2	0.3	0.1	0.2	0.1
Alkanes (C1-C4)	3.7	5.2	1.6		
Aliphatic aldehydes	7.4	10.1	3.8		
Unsaturated aliphatic hydrocarbons	2.9	4.1	1.2		
Fluorocarbons	1.2	1.8	0.3		
Acids and bases					
Hydrogen fluoride	0.9	1.3	0.3	1.6	2.5
Hydrogen sulphide	1.7	2.5	0.7		
Hydrogen chloride	3.8	4.7	2.5	4.2	4.0
Inorganic acid solutions	4.7	6.1	2.7	6.5	7.2
Caustic alkali solutions	2.7	3.3	1.9	3.0	2.0
Javel water	0.0	0.0	0.0		
Plating solutions	0.3	0.4	0.1	0.2	0.1
All acids	0.1	0.0	0.1		
Nitric acid	0.6	0.5	0.6	0.0	0.1
Hydrogen peroxide	0.5	0.3	0.6		
Phosphoric acid	0.4	0.5	0.3	0.7	0.7
Sulphuric acid	3.1	3.8	2.2	3.1	4.4
Acetic acid	0.9	0.8	1.0	0.3	0.7
Formic acid	0.1	0.1	0.1	0.0	0.1
Carbon disulphide	0.1	0.1	0.1		
Acrylonitrile	0.0	0.0	0.0		
Methyl methacrylate	0.1	0.0	0.3	0.0	0.0
Phenol	0.1	0.1	0.2	0.0	0.2
Turpentine	0.7	0.9	0.3	1.3	1.6
Volatile Organic Liquids	38.7	46.7	27.8		

Group/agent	Global prevalence in PROtEuS (%) All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	Prevalence in blue-collar jobs in comparison (%) PROtEuS[4]	Lung[5]
Cyanides	0.8	1.2	0.2		
Fluorides	1.0	1.4	0.4		
Hypochlorites	2.8	4.0	1.1	5.7	2.7
Nitrates	0.0	0.0	0.0	0.0	0.1
Other organic products					
Alkyds	1.3	2.0	0.4	2.7	3.2
Calcium oxide fumes	0.7	1.1	0.1	1.4	2.3
Glycerine	0.0	0.0	0.0		
Epichlorohydrin	0.0	0.0	0.0		
Nitroglycerine	0.0	0.0	0.0		
RDX	0.0	0.0	0.0		
Trinitrotoluene	0.0	0.0	0.0		
Polychlorinated biphenyls or PCBs	0.4	0.5	0.2	0.4	0.3
Aromatic amines	1.4	2.0	0.6	2.2	2.4
Phthalates	0.3	0.5	0.1	0.7	0.5
Isocyanates	0.6	1.0	0.2	1.0	0.4
Organic Sulfur Compounds	0.0	0.0	0.0		
Organic dusts					
Organic dyes and pigments	2.3	3.1	1.1	3.0	3.1
Cotton dust	3.7	5.7	1.1	4.9	5.1
Wool fibres	1.6	2.5	0.4	2.3	3.0
Silk fibres	0.4	0.6	0.1	0.8	1.1
Wood dust	9.2	12.6	4.5	15.1	19.1
Grain dust	1.2	1.9	0.3	2.7	5.3
Flour dust	1.2	1.9	0.3	2.3	2.8
Fur dust	0.2	0.3	0.0	0.1	0.1
Flax fibres	0.4	0.6	0.1	0.6	0.9
Cork dust	0.1	0.2	0.0	0.0	0.1
Hair dust	0.3	0.5	0.0		
Starch dust	1.6	2.5	0.3	2.1	2.0
Sugar dust	0.7	1.1	0.1	1.5	2.0
Rosin	0.1	0.0	0.1		
Felt dust	0.0	0.0	0.0	0.0	0.1
Leather dust	0.8	1.2	0.2	0.3	0.2
Tobacco dust	0.1	0.2	0.0	0.2	0.3
Natural rubber	0.2	0.3	0.0	0.0	0.2

Group/agent	Global prevalence in PROtEuS (%)			Prevalence in blue-collar jobs in comparison (%)	
	All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	PROtEuS[4]	Lung[5]
Tannic acid	0.1	0.1	0.0		
Synthetic fibres	3.1	4.6	1.0		
Plastic dusts	3.0	4.2	1.4	5.0	2.4
PVC dust	0.0	0.0	0.0		
Rayon fibres	0.7	1.1	0.2	1.0	1.1
Acrylic fibres	0.5	0.8	0.2	0.5	0.7
Polyester fibres	2.2	3.3	0.7	2.8	2.5
Nylon fibres	1.2	1.8	0.5	1.5	1.7
Acetate fibres	0.2	0.3	0.0	0.2	0.5
Cellulose acetate	0.1	0.2	0.0	0.2	0.3
Cellulose nitrate	0.3	0.5	0.0	0.4	0.2
Polyethylene	0.1	0.1	0.1		
Polypropylene	0.1	0.1	0.0		
Polystyrene	0.4	0.6	0.2	0.6	0.4
Poly(vinyl chloride)	0.5	0.7	0.2	0.4	0.5
Poly(vinyl acetate)	0.7	1.1	0.1	1.3	2.3
Polyamides	0.0	0.1	0.0	0.1	0.0
Polyacrylates	1.0	1.2	0.6	1.6	1.5
ABS (acrylonitrile-butadiene-styrene)	0.2	0.3	0.0	0.6	0.2
Epoxies	0.9	1.3	0.5	1.9	0.7
Phenol-formaldehyde	0.1	0.2	0.1	0.1	0.2
Urea-formaldehyde	0.6	0.9	0.2	1.6	0.2
Melamine-formaldehyde	0.1	0.2	0.1	0.2	0.1
Polyurethanes	1.1	1.5	0.5	1.9	1.3
Polyesters	0.5	0.8	0.1	1.5	0.6
Styrene-butadiene rubber	0.2	0.4	0.1	0.1	0.2
Polychloroprene	0.0	0.0	0.0		
Treated textile fibres	4.0	6.0	1.3	5.9	5.8
Untreated textile fibres	0.5	0.9	0.1	0.1	0.5
Cellulose	6.4	8.9	3.0	10.5	9.5
Rubber dust	1.1	1.6	0.4	1.5	1.2
PVC	0.0	0.0	0.0		
Inorganic dusts					
Abrasives dust	9.5	13.9	3.6	16.3	11.7
Inorganic insulation dust	5.2	7.5	2.2	7.3	6.8
Construction site dust	0.0	0.0	0.1		

Group/agent	Global prevalence in PROtEuS (%)			Prevalence in blue-collar jobs in comparison (%)	
	All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	PROtEuS[4]	Lung[5]
Soil dust	12.6	16.2	7.8	21.1	19.7
Chrysotile asbestos	6.8	10.4	2.0	10.1	10.5
Amphibole asbestos	1.6	2.5	0.4	2.2	3.2
Cristalline silica	7.2	9.4	4.1		
Portland cement	3.5	4.3	2.3	5.5	5.9
Glass dust	0.3	0.5	0.1	0.2	0.3
Glass fibres	2.0	3.0	0.5	3.3	2.4
Industrial talc	0.8	1.3	0.2	1.7	2.5
Brick dust	1.8	2.7	0.6	2.4	2.9
Clay dust	0.4	0.7	0.1	0.8	2.0
Concrete dust	6.6	8.8	3.7	9.6	8.4
Refractory brick dust	0.1	0.1	0.0	0.1	0.4
Mineral wool fibres	2.7	3.8	1.1	4.3	5.2
Extenders	2.5	3.7	0.9	4.9	5.0
Mica	0.1	0.2	0.0	0.0	0.4
Perlite, vermiculite	0.1	0.1	0.0	0.2	0.1
Cosmetic talc	1.1	0.9	1.4	1.2	0.3
Sodium carbonate	0.4	0.6	0.2	0.2	1.3
Sodium hydrosulphite	0.4	0.4	0.4		
Silicon carbide	3.4	5.4	0.6	7.9	5.9
Phosphorus	0.0	0.0	0.0		
Sulfur	0.2	0.3	0.1	0.1	0.5
Calcium oxide	1.9	2.9	0.4	3.0	3.9
Calcium sulphate	7.4	6.4	8.7	7.5	8.0
Calcium carbonate	8.1	5.7	11.4	7.3	9.0
Tungsten carbide	0.2	0.4	0.1	0.5	0.4
Silicates	0.0	0.0	0.0		
Radiation, electric and magnetic fields					
Ionizing radiation	1.0	0.6	1.6	0.3	0.4
Ultraviolet radiation	11.3	16.0	5.0	20.6	27.0
Pesticides					
DDT	0.2	0.2	0.1	0.2	0.7
Dichlorobenzene	0.0	0.0	0.0		
Pentachlorophenol	0.0	0.1	0.0		
Creosote	0.4	0.7	0.1	0.8	0.6
Pesticides	2.1	3.1	0.7		

Group/agent	Global prevalence in PROtEuS (%)			Prevalence in blue-collar jobs in comparison (%)	
	All jobs[1]	Blue collar jobs[2]	White collar jobs[3]	PROtEuS[4]	Lung[5]
Fungicides	0.1	0.1	0.0	0.2	0.0
Insecticides	1.8	2.7	0.5		
Herbicides	0.1	0.3	0.0	0.3	0.7
Wood preservatives	0.3	0.6	0.0	0.5	0.4
General categories					
Animal Fat	0.0	0.0	0.0		
Cleaning agents	12.6	19.3	3.6	22.7	17.8
Cosmetics	0.0	0.0	0.0		
Pharmaceuticals	0.2	0.2	0.4		
Photographic products	0.3	0.2	0.5		
Laboratory products	0.0	0.0	0.0		
Fertilizers	0.8	1.4	0.1	1.9	2.6
Biocides	5.3	6.9	3.3	8.9	10.6
Bleaches	2.1	3.0	0.9	4.0	0.7
Antineoplastic medication	0.0	0.0	0.0		
Microorganisms	8.2	11.3	4.1	16.0	14.7

1. Percentage of all jobs in PROtEuS (n=16,065) exposed by agent
2. Percentage of all blue collar jobs in PROtEuS (n=9239) exposed by agent
3. Percentage of all white collar jobs in PROtEuS (n=6826) exposed by agent
4. Percentage of jobs in PROtEuS retained in the comparison (n=4318) exposed by agent
5. Percentage of jobs in the Lung cancer study retained in the comparison (n=3022) exposed by agent

Supplementary table III. Cumulative odds ratios and 95% confidence intervals for the association between the categorical reliability, concentration or frequency of exposed and source of exposure data

	N Lung (1)	N PROtEuS (2)	Reliability OR (95% CI) (3)	Concentration OR (95% CI)	Frequency OR (95% CI)
Overall	18864	29551	**1.31 (1.26-1.36)**	**0.92 (0.88-0.95)**	**0.88 (0.85-0.91)**
Occupations with Expert-annotated JEPs (n=75)	18089	28210	**1.32 (1.27-1.37)**	**0.93 (0.89-0.97)**	**0.90 (0.87-0.93)**
Chemical group (4)					
Metals (n=19)	2181	4223	1.09 (0.99-1.21)	1.09 (0.98-1.21)	**1.13 (1.02-1.25)**
Organic Solvents (n=17)	1590	2782	**1.30 (1.14-1.47)**	**1.85 (1.63-2.09)**	**1.19 (1.06-1.34)**
Paints, varnishes and inks (n=8)	619	910	1.17 (0.93-1.48)	**0.80 (0.65-0.98)**	1.11 (0.92-1.34)
Combustion products and engine emissions (n=16)	4412	6101	**1.36 (1.26-1.46)**	**0.60 (0.54-0.66)**	**0.78 (0.73-0.85)**
Petroleum products (n=9)	839	1308	**1.35 (1.13-1.62)**	0.93 (0.74-1.17)	**0.56 (0.46-0.66)**
Complex organic fluids (n=7)	743	1353	**1.42 (1.17-1.73)**	**1.35 (1.08-1.69)**	0.99 (0.84-1.17)
Gases (n=7)	534	929	**1.58 (1.28-1.94)**	**0.50 (0.25-0.99)**	**0.58 (0.47-0.70)**
Acids and bases (n=9)	545	989	0.97 (0.80-1.18)	1.20 (0.90-1.60)	0.91 (0.74-1.11)
Organic dusts (n=31)	1890	2838	**1.25 (1.11-1.41)**	**0.85 (0.74-0.97)**	**0.88 (0.79-0.98)**
Inorganic dusts (n=24)	3109	4558	**1.21 (1.11-1.32)**	**0.62 (0.56-0.69)**	**1.13 (1.03-1.23)**
Radiation, electric and magnetic fields (n=2)	784	841	**2.78 (2.24-3.44)**	—	0.95 (0.76-1.19)
General categories (n=5)	1221	2053	**1.29 (1.12-1.49)**	**0.36 (0.30-0.43)**	**0.77 (0.67-0.89)**
By 2-digit CCDO (5)					
41: Clerical and Related Occupations (n=10)	535	716	**1.36 (1.09-1.69)**	**0.34 (0.18-0.64)**	**0.42 (0.33-0.52)**
51: Sales Occupations (n=10)	688	1496	**0.81 (0.68-0.96)**	**0.44 (0.29-0.67)**	**0.83 (0.70-0.99)**
61: Service Occupations (n=14)	2007	3461	**1.16 (1.05-1.30)**	**0.61 (0.52-0.73)**	**0.87 (0.78-0.96)**
71: Farming. Horticultural and Animal-Husbandry Occupations (n=5)	1088	778	**1.38 (1.15-1.65)**	**0.79 (0.65-0.96)**	**0.72 (0.61-0.86)**
83: Machining and Related Occupations (n=5)	1781	3370	**1.34 (1.20-1.50)**	1.00 (0.90-1.13)	0.96 (0.86-1.07)
85: Product Fabricating, Assembling and Repairing Occupations (n=9)	3398	6401	**1.20 (1.10-1.31)**	**1.22 (1.12-1.32)**	0.96 (0.89-1.04)
87: Construction Trades Occupations (n=12)	5233	7566	**1.18 (1.10-1.26)**	**0.74 (0.68-0.80)**	0.95 (0.89-1.01)

	N Lung (1)	N PROtEuS (2)	Reliability OR (95% CI) (3)	Concentration OR (95% CI)	Frequency OR (95% CI)
91: Transport Equipment Operating Occupations (n=10)	2010	2665	**1.93 (1.71-2.17)**	**0.29 (0.23-0.38)**	**0.78 (0.70-0.88)**
93: Material-Handling and Related Occupations, n.e.c. (n=7)	685	757	1.20 (0.99-1.46)	**0.44 (0.34-0.56)**	**1.59 (1.29-1.97)**
By time period (n) (6)					
1953-1993 (n=84)	15647	26532	**1.24 (1.19-1.29)**	**0.92 (0.88-0.96)**	**0.93 (0.89-0.96)**
1934-1972 (n=40) (7)	8775	9074	**1.29 (1.22-1.37)**	**0.90 (0.84-0.96)**	0.96 (0.91-1.01)
1973-2012 (n=40) (7)	6331	12824	**1.14 (1.08-1.21)**	**0.80 (0.75-0.86)**	1.02 (0.96-1.07)

1. Number of unique exposures (job/agent pairs) in the Lung study data
2. Number of unique exposures (job/agent pairs) in the PROtEuS data
3. Cumulative odds ratio (95% confidence interval)
4. Chemical/Physical groups with at least 50 concordant exposed occupation-agent combinations shown; n=number of agents within group
5. 2-digit CCDO groups with at least 50 concordant exposed occupation-agent combinations shown; n=number of 7-digit occupations within 2-digit group
6. n represents the number of 7-digit occupations included in the time period
7. Data restricted to the same set of 7-digit occupations in the two time periods

Supplementary table IV. Agreement in exposure status among occupation-agent combinations between Lung and PROtEuS exposure data

		Percent concordant		Percent discordant		Concordant exposed occupation-agent combinations		
	Number of occupation/ agent combinations (1)	Exposed (%)	Unexposed (%)	Exposed Lung (%) (2)	Exposed PROtEuS (%) (3)	Number of combinations (4)	Kendall correlation in probability (5)	Median diff. in probability, PROtEuS-Lung (6)
Overall	18270	8.2	85.5	4.1	2.2	1502	0.55	1.3
Occupations with Expert-annotated JEPs (n=75) (7)	15225	9.1	84.5	4.0	2.3	1393	0.55	1.5
By chemical group (8)								
Metals (n=21)	1890	10.3	83.8	4.3	1.7	194	0.60	0.1
Organic Solvents (n=17)	1530	9.3	83.1	5.0	2.6	142	0.59	1.2
Paints, varnishes and inks (n=8)	720	8.1	84.7	5.3	1.9	58	0.58	2.2
Combustion products and engine emissions (n=19)	1710	17.1	74.1	5.3	3.5	293	0.54	0.2
Petroleum products (n=11)	990	8.2	84.1	3.9	3.7	81	0.34	4.9
Complex organic fluids (n=8)	720	9.7	82.8	4.2	3.3	70	0.38	11.5
Gases (n=12)	1080	5.0	89.2	4.1	1.8	54	0.48	-1.2
Acids and bases (n=15)	1350	4.4	89.9	4.7	1.0	59	0.44	1.1
Organic dusts (n=41)	3690	4.7	90.3	3.4	1.6	174	0.60	0.0
Inorganic dusts (n=25)	2250	9.2	84.7	3.5	2.6	208	0.53	5.1
Radiation, electric and magnetic fields (n=2)	180	20.6	70.6	5.6	3.3	37	0.69	0.0
General categories (n=5)	450	21.3	63.3	8.2	7.1	96	0.55	2.4
By 2-digit CCDO major group (9)								
41: Clerical and Related (n=10)	2030	2.7	92.1	4.0	1.3	54	0.44	-5.3
51: Sales (n=10)	2030	4.3	90.6	2.1	3.0	88	0.43	10.5
61: Service (n=14)	2842	5.5	89.7	3.3	1.5	157	0.54	2.0
71: Farming. Horticultural and Animal-Husbandry (n=5)	1015	7.2	87.3	1.0	4.5	73	0.53	0.0
83: Machining and Related (n=5)	1015	15.0	75.1	7.3	2.7	152	0.63	-0.6
85: Product Fabricating, Assembling and Repairing (n=9)	1827	15.3	76.0	6.3	2.4	279	0.57	0.9
87: Construction Trades (n=12)	2436	17.0	74.7	6.1	2.2	414	0.54	3.7
91: Transport Equipment Operating (n=10)	2030	5.0	91.1	2.4	1.6	101	0.62	-2.2
93: Material-Handling and Related, n.e.c. (n=7)	1421	6.0	84.2	6.8	3.0	85	0.53	-1.3

	Number of occupation/ agent combinations (1)	Percent concordant		Percent discordant		Concordant exposed occupation-agent combinations		
		Exposed (%)	Unexposed (%)	Exposed Lung (%) (2)	Exposed PROtEuS (%) (3)	Number of combinations (4)	Kendall correlation in probability (5)	Median diff. in probability, PROtEuS -Lung (6)
By time period (10)								
1953-1993 (n=84)	17052	8.2	85.3	4.3	2.2	1400	0.56	1.4
1934-1972 (n=40) (e)	6880	8.8	83.8	4.2	3.2	604	0.58	0.4
1973-2012 (n=40) (e)	6880	9.2	82.1	5.8	3.0	631	0.58	1.9

1. Number of occupation-agent combinations with at least 10 jobs in both PROtEuS and Lung
2. Proportion of occupation-agent combinations with probability ≥5% in Lung and <5% in PROtEuS
3. Proportion of occupation-agent combinations with probability ≥5% in PROtEuS and <5% in Lung
4. Number of concordant exposed occupation-agent combinations (probability ≥5% in Lung and ≥5% in PROtEuS)
5. Kendall correlation in the probability of exposure between concordant exposed occupation-agent combinations
6. Median difference in probability (probability in PROtEuS minus probability in Lung) across concordant exposed occupation-agent combinations
7. Restricted to combinations with occupations with reviewed JEP (75 7-digit CCDO)
8. Chemical/Physical groups with at least 50 concordant exposed occupation-agent combinations shown; n=number of agents within group
9. 2-digit CCDO groups with at least 50 concordant exposed occupation-agent combinations shown; n=number of 7-digit occupations within 2-digit group
10. n represents the number of 7-digit occupations included in the time period
11. Data restricted to the same set of 7-digit occupations in the two time periods

Chapitre 7. Discussion générale

Discussion générale

La caractérisation de l'exposition professionnelle rétrospective représente un défi dans les études de population, car elle implique d'évaluer un grand nombre d'emplois répartis entre une multitude de milieux de travail sur une longue période temporelle. Le manque d'information objective spécifique à chaque emploi constitue une difficulté majeure à la capacité d'offrir des estimations justes et précises. Ce travail a porté sur la valorisation de la banque de données d'expertises montréalaises, associées à une approche d'évaluation considérée comme la référence dans le domaine, pour développer une MEE représentant une source d'information sur l'exposition rétrospective adaptée à une diversité d'applications et de contextes. Les recherches présentées ont également permis d'aborder certaines limites de cette MEE, en développant notamment un cadre de modélisation pour raffiner les estimations pour les cellules basées sur un faible nombre d'emplois. Une autre limite, formée par l'utilisation de catégories pour représenter l'intensité de l'exposition dans les cellules, a été abordée en combinant CANJEM à une banque de mesures historiques. Cet effort a permis d'obtenir des niveaux d'exposition quantitatifs pour toutes les cellules de CANJEM, documentées ou non dans la banque de données de mesures. Ce travail a finalement permis de caractériser l'impact d'une nouvelle approche hybride d'évaluation, fondée sur une compilation de données existantes, sur les expositions attribuées par une comparaison avec les expositions assignées par expertise traditionnelle dans une étude antérieure.

Ce chapitre présente un survol des principaux points soulevés pour chacun des volets. Certaines limites associées à l'utilisation de la banque d'expertises montréalaises sous-tendant l'ensemble des travaux sont également discutées. L'originalité du travail, et le potentiel d'application et de bonification des résultats afin de poursuivre l'objectif d'amélioration de l'évaluation de l'exposition professionnelle dans les études de population et de la prévention des risques à la santé viennent clore ce chapitre.

7.1 Contribution à la recherche

7.1.1 Développement de la MEE CANJEM

Le développement de la matrice CANJEM a permis de constituer une source d'information rétrospective multiagents et multi-industries sur l'exposition à partir d'une synthèse d'expertises individuelles tirées d'études passées. La synthèse d'expertises individuelles appliquée dans le développement de CANJEM, par opposition à une estimation unique par profession propre aux MEE traditionnelles, a permis de refléter la variabilité de l'exposition entre les sujets à l'intérieur d'une profession ou industrie donnée. Cette hétérogénéité est exprimée par l'utilisation de la probabilité sous forme continue comme indicateur de la présence d'exposition, et de la distribution relative des emplois exposés par catégorie d'intensité plutôt qu'une seule valeur. L'application des informations d'une cellule de CANJEM pour attribuer un statut exposé/non-exposé et/ou un niveau d'intensité à un sujet dans une population d'étude requiert de réduire cette information en une seule valeur, par exemple en utilisant la catégorie d'intensité la plus fréquente. L'information sur la variabilité de l'exposition dans les cellules permet toutefois de choisir plusieurs paramètres pour représenter l'exposition et d'évaluer la sensibilité des analyses à ces choix. L'information sur la variabilité intra-profession inhérente à CANJEM peut être plus directement appliquée à d'autres utilisations, comme l'estimation du nombre de travailleurs exposés à différents niveaux dans la population pour une étude sur le fardeau des maladies professionnelles.

En outre, l'utilisation d'emplois individuels comme source de données pourrait également permettre de moduler les indices d'exposition des cellules en fonction des connaissances sur les mécanismes d'action des agents. À titre d'exemple, les données chez le rat suggèrent que le risque de cancer du poumon associée à une exposition à de faibles concentrations en fumées d'échappement de moteur diésel est faible puisque le mécanisme de clairance pulmonaire permet d'éliminer les particules inhalées, tandis qu'une exposition à une forte concentration dépassant la capacité de clairance et induit une réponse inflammatoire et des dommages aux tissus (Centre international de recherche sur le cancer, 2014). En conséquence, il pourrait être possible de calculer des indices d'exposition alternatifs en variant la pondération des catégories d'intensité dans le calcul de l'intensité d'exposition moyenne pondérée (p.ex. en augmentant les contrastes), ou en modifiant le seuil minimum d'intensité définissant un emploi exposé dans le calcul de la probabilité, afin d'explorer le risque de maladie associée à un agent en tenant compte des données expérimentales.

L'utilisation d'expertises individuelles passées dans la réalisation de CANJEM signifie toutefois que le nombre de données disponibles permettant de décrire le portrait de l'exposition peut varier d'une cellule à une autre. L'augmentation du niveau de résolution dans les professions, industries et périodes temporelles implique également de répartir un nombre d'emplois fixes entre un plus grand nombre de catégories. Une comparaison avec les données de deux recensements menés à 25 ans d'intervalle a permis d'observer une bonne couverture des données par rapport à la population générale, et ce, même à un niveau de résolution relativement détaillé. Par exemple, plus de 75% des emplois dans les populations montréalaise et canadienne dénombrés en 1986 ou en 2011 étaient associés à des groupes de professions (codes CCDP et CNP à 4 chiffres) pour lesquelles au moins 10 emplois étaient disponibles conjointement dans les périodes 1930-1969 et 1970-2005. La diminution du niveau de précision pour les professions ou les périodes, de la taille d'échantillon minimale par cellule, ou d'une combinaison de ces deux facteurs, permettaient d'augmenter le niveau de couverture de la population.

7.1.2 Modélisation des évaluations par expertise

Les travaux présentés au Chapitre 4 visaient à raffiner les estimations de l'exposition pour les cellules moins bien documentées, en développant une approche de modélisation permettant de partager l'information disponible entre les cellules de professions similaires. Par cette approche, l'estimation d'une cellule donnée représente un compromis entre la moyenne pour la profession, non biaisée, mais potentiellement imprécise, et la moyenne du groupe de professions (catégorie supérieure dans la hiérarchie), qui est quant à elle plus précise, mais moins spécifique. Pour une profession basée sur peu d'emplois, la faible précision de l'estimation entraînera une influence plus forte de la moyenne du groupe, comparativement à une profession basée sur un nombre plus grand d'emplois. Ce phénomène associé au partage d'information est utile lorsqu'il y a une cohérence entre les professions individuelles d'un même groupe.

L'attention particulière portée à l'effet de rétrécissement dans ce travail, rarement documenté dans la littérature sur l'évaluation de l'exposition, était motivée par l'utilisation d'une classification professionnelle standardisée pour structurer le partage d'information sur l'exposition entre les professions. Le regroupement des professions dans les classifications est principalement fondé sur des similitudes dans la nature des tâches, services et biens produits, qui peuvent être dans certains cas sans relation directe avec le profil d'exposition. Par exemple, la profession de pompiste ou commis de station-service est nichée dans le grand groupe des commis-vendeurs de biens de consommation. L'exposition à des vapeurs d'essence pour cette profession pourrait être plus apparentée à celle de professions du domaine de transport automobile comparativement aux autres types de commis-vendeurs tels les vendeurs de fournitures de jardins ou de pelouse ou de tissus à la verge. La création d'un regroupement des professions et industries basé sur les similitudes dans l'exposition plutôt que sur la nature du travail nécessiterait toutefois un apport d'expertise important qui devrait également prendre en considération l'ensemble des 258 agents dans CANJEM, et être réalisé pour chaque système de classification, ce qui est irréaliste. Puisque le compris entre l'estimation spécifique d'une profession individuelle et la moyenne du groupe n'est probablement pas systématiquement avantageux, nous avons cherché à limiter son application en évaluant l'ampleur du phénomène de rétrécissement pour des tailles d'échantillon différentes.

L'évaluation de l'amplitude du rétrécissement sur les estimations a permis d'observer des effets importants pour les cellules basées sur 4 emplois ou moins, modérés entre 5 et 9 emplois, puis négligeables au-dessus de 10 emplois. Une taille d'échantillon minimale de 5 emplois représente un compromis permettant de fournir une estimation plus précise sur l'exposition en utilisant l'information disponible dans les cellules de professions similaires, sans toutefois être allouer un effet de rétrécissement extrême. Ce seuil, bien qu'arbitraire, permet ainsi de limiter les impacts potentiels associés à une incompatibilité entre la structure des systèmes de classification standards et le profil d'exposition dans certaines situations, sachant que l'identification de telles situations à l'avance est impossible en raison de l'ampleur de CANJEM.

De manière globale, les résultats observés lors de l'application de la modélisation pour l'indice de l'intensité d'exposition moyenne pondérée étaient en accord avec les tendances attendues par l'utilisation de modèles hiérarchiques, soit des estimations plus extrêmes tirées vers la moyenne.

Les résultats pour la probabilité d'exposition étaient quant à eux plus mitigés puisqu'une sous-estimation systématique, bien que d'une amplitude modérée, a été observée pour le niveau des sous-groupes (codes à 3 chiffres), et peut résulter d'un manque de compatibilité des modèles de régression logistique dans un contexte où plus de 75% des cellules n'avaient aucun emploi exposé. À ce titre, des approches pour les distributions gonflées à zéro (« zero-inflated ») ont été développées pour des modèles de régression de Poisson et de régression logistique dans un cadre hiérarchique (Hall, 2000) et pourraient représenter une avenue possible pour la probabilité d'exposition. L'application de ce type de modèle à une structure hiérarchique comportant plusieurs niveaux est cependant complexe et sortait du cadre de ce travail. Elle représente néanmoins une perspective de recherche intéressante pour raffiner les estimations de la probabilité d'exposition dans les cellules de CANJEM.

7.1.3 Estimation de niveaux quantitatifs

La modélisation des mesures historiques contenues dans la banque CWED au Chapitre 5 a permis d'associer des niveaux quantitatifs pour l'exposition aux poussières de bois aux cellules de CANJEM avec une probabilité d'exposition non nulle, même pour des professions sans mesures spécifiques grâce à l'estimation de concentrations moyennes pour les catégories d'intensité faible, moyenne et élevée. À titre d'exemple, les moyennes géométriques prédites pour l'année 1989 correspondant à une cellule avec 100% des emplois exposés à une intensité faible, moyenne ou élevée étaient respectivement 0,75 mg/m^3, 1,02 mg/m^{3}et 1,68 mg/m^3. Pour une MEE traditionnelle où chaque cellule n'est associée qu'à une seule catégorie semi-quantitative, le niveau quantitatif attribué serait l'une ou l'autre de ces trois valeurs en l'absence de mesures spécifiques. Dans le cas de CANJEM, l'intensité dans une cellule est représentée par une distribution relative des niveaux semi-quantitatifs, ce qui permet d'attribuer des expositions sur un continuum entre 0,75 mg/m^3 (100% faible) et 1,68 mg/m^3 (100% élevé). L'information sur la variabilité des niveaux d'exposition des emplois permet ainsi d'estimer des concentrations plus contrastées entre les professions, même celles sans mesures associées.

Les valeurs prédites rapportées ci-dessus correspondent à un ratio de 1-1,3-2,3 dans les moyennes géométriques entre les catégories d'intensité. Ces contrastes sont comparables à ceux rapportés dans la littérature lors d'exercices similaires combinant une MEE générale à des mesures historiques (Tableau I ci-dessous). Par contre, ils sont plus faibles comparativement à l'échelle de 1-5-25 entre les niveaux faible, moyen et élevé utilisée pour le calcul de l'intensité d'exposition moyenne pondérée (abordé au Chapitre 3). Cette échelle représentait une estimation moyenne des ratios entre les niveaux d'intensité relatifs à travers l'ensemble des agents évalués dans les études montréalaises, établie à partir de discussion avec les experts. D'après eux, les contrastes entre les catégories d'intensité variaient entre les agents, pouvant aller de différences plus faibles (p.ex. échelle de 1-2-3) à des contrastes plus grands (p.ex. 1-10-100). L'échelle 1-5-25 représentait une valeur moyenne plausible en l'absence de ressource pour obtenir des contrastes spécifiques à chacun des 258 agents. Les résultats observés suggèrent donc que des contrastes plus faibles entre les catégories seraient plus adaptés pour les poussières de bois. La caractérisation des associations entre les catégories d'intensité et les niveaux quantitatifs de l'exposition pourrait également être approfondie par l'élargissement du champ d'application des modèles développés. Outre leur application à d'autres agents bien représentés dans CWED tels le formaldéhyde et la silice, le croisement avec d'autres banques de données d'exposition professionnelles permettrait d'évaluer les différences dans les associations observées entre les sources de mesures pour un même agent. Le croisement des sources de données sur la base des industries plutôt que sur les professions, et l'estimation des relations entre les niveaux d'exposition et l'indice de l'intensité d'exposition moyenne pondérée calculé selon différentes échelles de pondération représentent d'autres possibilités pour raffiner les tendances observées dans ce travail.

Tableau I. Ratios entre les moyennes géométriques estimées par catégorie d'intensité d'exposition recensés dans les études combinant une MEE générique et des mesures historiques d'hygiène industrielle

Étude	Agent	Point de comparaison	Ratio MG[1]
Peters et coll. (2016)	Benzo[a]pyrène	Élevé vs. faible	1.67
	Amiante	Élevé vs. faible	1.20
	Chrome hexavalent	Élevé vs. faible	1.20
	Nickel	Élevé vs. faible	0.69[2]
Peters et coll. (2011b)	Silice cristalline	Élevé vs. non-exposé	1.59
		Élevé vs. faible	1.64
Friesen et coll. (2012)	Benzène	Moyen vs. faible	1.60
		Élevé vs. faible	3.00
Koh et coll. (2014)	Plomb (fumées)	Élevé vs. faible	7.14
		Élevé vs. moyen	1.39
	Plomb (poussières)	Élevé vs. faible	2.63
		Élevé vs. moyen	1.96

1. Ratio entre la moyenne géométrique d'une catégorie relativement à la moyenne géométrique pour le niveau de références
2. La catégorie « faible » était associée à des niveaux moyens supérieurs à ceux pour la catégorie « élevée »

Par ailleurs, les mesures consignées dans les banques de données peuvent être associées à différents types de biais pouvant influencer l'interprétation des niveaux d'exposition, tels la sélection des milieux de travail évalués, la stratégie d'échantillonnage (p.ex. sélection des tâches ou postes de travail les plus exposés, vérification de la conformité réglementaire de l'entreprise) et l'enregistrement sélectif des résultats analytiques (Olsen, 1988; Sarazin et coll., 2016). L'étendue de la représentativité des mesures disponibles dans une banque pour caractériser l'exposition professionnelle à l'échelle de la population peut donc être difficile à établir en raison de ces facteurs. Ces biais peuvent ainsi avoir eu une incidence sur les niveaux d'exposition quantitatifs estimés pour CANJEM dans ce travail. Par exemple, les mesures disponibles pour les professions avec un niveau d'intensité principalement faible (p.ex. adjoints administratifs) pourraient surreprésenter des circonstances (p.ex. tâches, postes de travail) peu fréquentes de forte exposition, tandis qu'elles pourraient être plus représentatives de l'exposition usuelle pour les professions avec un niveau d'intensité élevé, causant ainsi un affaiblissement des contrastes entre les catégories d'intensité. Finalement, moins de la moitié des professions avec une probabilité d'exposition non nulle aux poussières de bois dans CANJEM avaient au moins une mesure disponible dans CWED. L'utilisation d'une banque de données d'exposition comme unique source d'information pour dresser un portrait des professions ou industries présentant un potentiel d'exposition pourrait ainsi mener à une sous-estimation de la situation réelle, montrant ainsi l'utilité des MEE comme source d'information complémentaire dans la caractérisation de l'exposition à l'échelle de la population.

7.1.4 Approche hybride d'évaluation de l'exposition

Les travaux présentés au Chapitre 6 visaient d'une part à décrire l'approche hybride utilisée dans l'étude PROtEuS, et d'autre part à évaluer l'impact de cette approche sur le nombre d'expositions par emploi et sur les niveaux de fiabilité, d'intensité et de fréquence assignés pour un sous-ensemble d'emplois de cols bleus. Les évaluations provenant de l'étude sur le cancer de poumon, réalisées selon une approche par expertise traditionnelle, représentaient quant à elles le groupe de référence dans les comparaisons.

Globalement, la distribution du nombre d'expositions assignées par emploi était comparable entre les deux approches. Par contre, les expositions assignées à l'aide de l'approche hybride étaient associées à un niveau de fiabilité plus élevé dans le jugement des experts quant à la présence de l'exposition, et cette tendance était constante dans les analyses stratifiées par groupe d'agent ou de profession et par période temporelle. Cette augmentation pourrait être expliquée par la plus grande quantité d'information sur l'exposition disponible systématiquement pour les experts afin de guider leur évaluation. Les niveaux d'intensité et de fréquence assignées avec l'approche hybride étaient sensiblement plus faibles comparativement à celles issues de l'étude sur le cancer du poumon, avec toutefois une différence de moindre amplitude. Les tendances observées étaient également variables dans les analyses stratifiées par groupe d'agent et de professions, et pourraient être dues en partie à des différences entre les deux études dans la répartition des emplois entre les 90 professions de cols bleus.

Les experts ont également eu tendance à assigner plus fréquemment des expositions à une même catégorie de fiabilité, intensité et fréquence au sein d'une combinaison de profession et d'agent, comparativement à l'approche traditionnelle. Cette observation peut découler de l'utilisation des informations dans les profils d'emplois, notamment des lignes directrices présentes dans le champ des commentaires, comme source de référence. Cette plus grande homogénéité observée dans la distribution des expositions peut être interprétée de deux façons différentes. D'un côté, elle peut représenter une plus grande cohérence dans les expositions assignées à des emplois similaires. D'un autre côté, elle pourrait aussi être le résultat d'une tendance à appliquer directement l'information du profil aux emplois, sans considérer les particularités propres à chaque sujet, se rapprochant d'une approche analogue à une MEE traditionnelle. L'amplitude du phénomène était cependant modérée dans nos analyses, et les expositions assignées demeuraient variables à l'intérieur des professions traduisant la prise en compte de différences individuelles dans les évaluations réalisées par cette approche.

7.2 Validité des estimations de l’exposition

Les travaux réalisés dans cette thèse ont tous comme point de départ une banque de données d’évaluations réalisées selon une approche par expertise, dont la synthèse a permis de fournir un portrait de l’exposition à de multiples agents par profession ou industrie. La validité des estimations de l’exposition dans les cellules de CANJEM ou dans les profils d’emploi de PROtEuS, et leur utilité comme source d’information sur l’exposition dans la population dépend de plusieurs facteurs. Un premier élément concerne la qualité du jugement porté sur chaque emploi, c'est-à-dire la capacité des experts à bien caractériser l’exposition. Un second élément concerne la validité des estimations synthétisées au sein d’une cellule et leur applicabilité à la population source des études, soit celle de la région métropolitaine montréalaise. La synthèse des expertises individuelles pour calculer chaque cellule a impliqué la mise en commun des informations provenant d’un échantillon non aléatoire de la population, puisque les données proviennent d’études cas-témoins. Ainsi, même si l’exposition de chaque sujet était parfaitement caractérisée, l’estimation d’une cellule pourrait donner un portrait potentiellement biaisé de l’exposition dans la population. Par exemple, si l’exposition à un facteur de risque fortement associé à une maladie était plus élevée chez les cas, et que ceux-ci représentaient la moitié des sujets dans une cellule, celle-ci pourrait fournir une surestimation de l’exposition dans la population générale. Finalement, même si l’information représente un portrait valide de l'exposition dans la population montréalaise, la validité des estimations pour caractériser l'exposition dans une autre population (p.ex. reste du Canada) pourrait être limitée.

Ces problématiques peuvent être mises en parallèle avec les questionnements entourant la création d’un portrait de l’exposition basé sur les mesures d’une banque de données. En premier lieu, la validité des mesures individuelles pour caractériser l'exposition dépendrait de la performance de la méthode de prélèvement et d'analyse de chaque échantillon, au même titre que la qualité des évaluations de experts. Ensuite, la validité des estimations par profession dépendrait notamment de la stratégie d'échantillonnage et de la sélection de milieux de travail visités, qui pourraient mener à une surestimation (par un ciblage préférentiel des situations à risque) ou à une sous-estimation (absence de mesures pour une profession ou industrie avec un potentiel d'exposition) relativement à la population. Finalement, la représentativité des estimations sur l'exposition pour une autre population pourrait être influencée par des différences dans l'environnement réglementaire ou dans les secteurs industriels.

7.2.1 Validité des expertises individuelles

L'évaluation de la validité des expositions assignées par les experts est limitée par l'absence d'un « étalon-or » servant de référence (McGuire et coll., 1998). Par ailleurs, le croisement entre les estimations d'intensité des cellules de CANJEM et les mesures d'exposition de la banque CWED ne constituait pas un exercice de validation en ce sens puisque les mesures ne provenaient pas des mêmes lieux de travail des sujets évalués. Néanmoins, cette analyse a permis d'observer un accord dans les concentrations moyennes par catégorie d'intensité, qui fut généralement constant dans les analyses de sensibilité.

L'approche par expertise, telle que développée et appliquée dans les études montréalaises, a été considérée comme la méthode de référence pour évaluer l'exposition rétrospective en l'absence de mesures objectives pertinentes puisqu'elle permet d'assigner des niveaux spécifiques à chaque description d'emploi, bien qu'elle soit complexe à mettre en œuvre (Bouyer et Hémon, 1993). Par contre, une revue des études de validité par Teschke et coll. (2002) a montré une performance variable de la méthode par expertise appliquée par différents groupes de recherche comparativement aux expositions auto-rapportées par les sujets ou à des mesures environnementales ou biologiques. Cette performance peut être améliorée par une évaluation de la fiabilité et de la validité des expertises dans l'étude et par une collecte d'information détaillée et variée pour documenter l'exposition. Les évaluations menées au sein du groupe montréalais ont d'ailleurs montré une bonne fiabilité et une bonne cohérence inter-experts dans les estimations (Goldberg et coll., 1986; Siemiatycki et coll., 1997), incluant celles assignées par une approche hybride analogue à celle utilisée pour l'étude PROtEuS (Kachuri et coll., 2016). D'autres évaluations ont permis de montrer la validité des historiques d'emploi rapportés par les sujets (Baumgarten et coll., 1983) et la capacité des experts montréalais à identifier correctement la présence d'expositions à divers agents dans l'environnement de travail à partir de descriptions d'emploi relativement sommaires (Fritschi et coll., 2003).

7.2.2 Validité des estimations pour la population montréalaise

Les recherches menées dans cette thèse ont porté sur les données d'emploi et d'exposition professionnelle tirées de cinq études menées sur une plage de plus de 30 ans, où chaque population d'étude a été définie au départ par le recrutement de cas d'un ou plusieurs sites de cancer. Les témoins pouvaient quant à eux provenir de la population générale ou de sujets atteints d'autres cancers. Le recrutement pouvait être limité aux hommes ou aux femmes, ou contenir les deux. De plus, les entrevues pouvaient avoir été menées avec un mandataire lorsque le sujet en était incapable, dans une proportion plus élevée dans le cas de maladies plus invalidantes (p.ex. approximativement 20% dans l'étude poumon). Finalement, la méthode d'évaluation de l'exposition a évolué sensiblement au cours des études. Ces différents facteurs auraient pu influencer la représentativité des estimations dans les cellules de CANJEM les profils d'emploi de l'approche hybride relativement à la population source des études, soit celle de la grande région de Montréal.

L'identification de différences systématiques dans les niveaux d'exposition des emplois tenus entre les cas et les témoins, et entre les hommes et les femmes, a fait l'objet de deux études publiées. Une étude a été menée par Kirkham et coll. (2016) à partir des données de l'étude poumon, dans laquelle les données d'emploi et d'exposition ont été agglomérées par profession séparément pour les cas et les témoins. Un accord élevé a été observé dans la probabilité d'exposition et l'intensité d'exposition moyenne pondérée par profession, et ce, même pour les agents figurant parmi la liste de cancérogènes avérés (Groupe 1 du CIRC) pour le site du poumon pour lesquels une différence systématique aurait pu être attendue. Une bonne concordance dans les indices de probabilité et d'intensité d'exposition moyenne pondérée entre les emplois des hommes et des femmes pour une même profession a également été observée (Labrèche et coll., 2014). Ces études n'ont pas montré d'incompatibilités majeures dans les niveaux d'exposition empêchant la mise en commun des données.

Pour les différences reliées au statut de répondant, les historiques professionnels obtenus de mandataires concernaient une plus grande proportion des sujets pour les cancers plus invalidants, représentant approximativement 20% des sujets inclus dans les études 1 et 2, comparativement à 9% pour l'étude 3 (cancer du sein) et 3% pour l'étude PROtEuS. L'historique professionnel provenant de mandataires comportait en moyenne moins d'emplois que ceux rapportés par les sujets. Les experts avaient tendance à assigner des expositions avec un degré de fiabilité plus faible, en raison notamment des descriptions d'emploi moins détaillées. Les niveaux d'intensité et de fréquence d'exposition assignées à ces emplois étaient toutefois comparables à ceux attribués aux emplois rapportés par les sujets.

Certains facteurs pouvant mener à des différences dans les expositions assignées pour des situations similaires entre les études incluent une expérience accrue avec la méthode d'évaluation, des modifications aux niveaux représentatifs des catégories d'intensité, des changements dans la composition de l'équipe d'experts, l'informatisation du processus d'encodage et l'emploi de questionnaires plus détaillés. L'évaluation de ces différences, plus subtiles, et leur ajustement sur les estimations de CANJEM, ont représenté une des facettes des analyses au Chapitre 4. L'analyse portant sur cinq agents n'a pas permis toutefois de mettre en évidence des différences systématiques entre les études sur la probabilité d'exposition et les niveaux d'intensité d'exposition moyenne pondérée des emplois. Les différences observées peuvent par ailleurs également représenter des différences résiduelles entre les sujets qui n'ont pas entièrement été prises en compte par la profession ou la période temporelle. L'utilisation de la modélisation statistique a également permis de pondérer l'influence d'une étude ou d'une autre sur les estimations des cellules afin de tenir compte de ces différences.

Finalement, les experts avaient accès à une compilation sommaire par profession et par agent des expertises réalisées précédemment dans l'étude multisite comme source d'information dans le codage des études poumon, sein et cerveau. Ainsi, bien que chaque emploi soit indépendant des autres, les expositions assignées étaient partiellement basées sur les évaluations passées, dans une bien moindre mesure toutefois que dans l'approche par expertise hybride puisque la documentation (sous forme de longs tableaux imprimés et difficiles à naviguer) n'était pas utilisée de manière systématique. Cette dépendance partielle pourrait causer une certaine sous-estimation de l'incertitude puisque le nombre d'emplois dans une cellule ne représenterait pas nécessairement un échantillon d'observations réellement indépendantes.

7.2.3 Validité des estimations pour d'autres populations

La déclinaison de CANJEM en de multiples systèmes de classification vise à faciliter son utilisation par la communauté de chercheurs internationaux et à combler un manque d'information sur l'exposition professionnelle rétrospective en Amérique du Nord et dans le monde. Or, bien que les estimations de CANJEM soient bien représentatives du contexte montréalais où les industries du textile, du vêtement, de l'alimentation et du transport ont historiquement représenté des moteurs économiques importants (Siemiatycki et Richardson, 1991; Brodeur et Galarneau, 1994) son application dans un cadre plus large doit prendre en compte les différences potentielles dans les industries ou le cadre législatif entre les populations. Par exemple, la pertinence des estimations de CANJEM pourrait être limitée pour des pays en voie de développement dont l'économie repose principalement sur le secteur primaire (agriculture, ressources naturelles), composée d'une grande part de petites entreprises et d'entreprises « informelles » et dont la règlementation en santé et sécurité du travail demeure limitée et peu appliquée (Hogstedt et Pieris, 2000; Giuffrida et coll., 2002; Pingle, 2012; Mrema et coll., 2015). Notons toutefois que CANJEM pourrait représenter une source d'information utile pour caractériser l'exposition pour des industries et procédés qui furent jadis importants à Montréal mais ayant fait l'objet de délocalisation vers l'étranger, tels les tanneries (Sahasranaman, 2000) et la production de vêtements (Ministères des finances du Québec, 2005; Hobson, 2013).

À l'échelle canadienne, la couverture par CANJEM de la population active dénombrée lors de deux recensements était comparable à celle observée pour la région montréalaise. La combinaison de CANJEM avec les mesures de la banque CWED, provenant de l'Ontario et de la Colombie-Britannique, a quant à elle permis une comparaison sur l'exposition à un agent spécifique. Sur 78 professions avec au moins une mesure, 70 étaient associées à une probabilité d'exposition non nulle dans CANJEM, tandis qu'une association positive entre l'augmentation des niveaux semi-quantitatifs de l'intensité dans les cellules et les concentrations a également été observée. Pour les États-Unis, CANJEM a été récemment utilisée pour estimer l'exposition à des pesticides dans une étude menée au Connecticut (Zeng et coll., 2017). Bien que la représentativité de CANJEM à la population de cet état de la Nouvelle-Angleterre n'ait pas fait l'objet d'une évaluation détaillée, la proximité économique de ces deux régions font en sorte qu'elles partagent plusieurs traits communs et constituent a priori des contextes comparables. À ce titre, la banque IMIS pourrait constituer une source d'information intéressante pour évaluer la pertinence des estimations de CANJEM à l'échelle des États-Unis et pour des régions administratives spécifiques.

Finalement, une comparaison entre les données d'exposition de l'étude poumon et les estimations de la matrice FINJEM portant sur 27 agents a été menée par Lavoué et coll. (2012b). Cette matrice, développée pour la population finlandaise, a depuis été appliquée – parfois directement – à d'autres populations (p.ex. Espagne, Australie, Pays-Bas) (Kauppinen et coll. 2014), et la comparaison aux données montréalaises visait à adapter les estimations pour une étude internationale sur le cancer du cerveau (van Tongeren et coll., 2013). Les niveaux de probabilité et d'intensité d'exposition par profession étaient comparables pour la majorité des agents évalués, bien que des différences aient été observées entre autres pour les poussières de farine et le toluène (Lavoué et coll., 2012c). Les banques de données, les données de la littérature et l'expertise locale pourraient être utilisées pour réaliser une évaluation de la compatibilité des estimations de CANJEM à une autre population. Une telle évaluation doit toutefois tenir compte des limites des approches dans l'interprétation de différences, qui peuvent provenir de véritables différences dans le profil d'exposition, ou de différences provenant des méthodes d'évaluation de l'exposition elles-mêmes.

7.3 Originalité de la recherche

Cette recherche est axée sur l'exploitation d'une banque de données d'expertises passées à des fins de source d'information sur l'exposition professionnelle rétrospective à une liste exhaustive d'agents, dont plusieurs ont une très faible prévalence d'exposition dans la population et sont rarement représentés dans les sources actuelles. De plus, l'utilisation de données existantes dans les études de population reste relativement peu fréquente et ont principalement porté sur des objectifs méthodologiques pour comparer des méthodes d'évaluation ou le portrait d'exposition entre des sous-populations (Dewar et coll., 1991; Peters et coll., 2011a; Lavoué et coll., 2012c; Labrèche et coll., 2014; Kirkham et coll., 2016).

L'utilisation d'emplois individuels dans le développement de CANJEM a permis de structurer l'information l'exposition selon plusieurs systèmes de classification, dont chacun était décliné en différents niveaux de résolution, et de décrire le profil d'exposition pour chaque cellule sous différents paramètres en tenant compte de l'hétérogénéité entre les emplois. CANJEM représente une des rares MEE, avec les matrices du programme MATGÉNÉ (Févotte et coll., 2011), qui ont été définies au départ en de multiples classifications professionnelles et industrielles.

La recherche a également permis d'entreprendre le développement d'un cadre méthodologique permettant une synthèse plus sophistiquée des expertises individuelles, comparativement à une approche descriptive, afin d'améliorer les estimations à travers deux dimensions d'une MEE. L'application de modèles hiérarchiques pour permettre le partage d'information entre les cellules de professions apparentées demeure une approche relativement peu utilisée dans le cadre d'études de population. Les modèles permettant d'assigner des niveaux quantitatifs à une MEE générales par une combinaison avec une banque de mesures ont quant à eux été basés sur une approche existante, qui a toutefois été adaptée pour tenir compte du format différent de l'intensité dans les cellules, soit une distribution plutôt qu'une valeur unique. Ce travail a par ailleurs représenté une des premières applications de la banque CWED dans une étude de population.

Finalement, la comparaison des expositions assignées dans l'étude PROtEuS avec celles issues d'une étude codée avec une approche par expertise traditionnelle a permis de réaliser un premier portrait de certains impacts associés à l'approche par expertise hybride, fondée sur une synthèse d'évaluations passées comme source d'information. Cette approche figure parmi un groupe de méthodes récemment développées, telles les règles de décision (Fritschi et coll., 2009; Friesen et coll., 2013), qui se situent à un point intermédiaire dans le continuum entre l'assignation purement individuelle de l'approche par expertise traditionnelle, et l'assignation par groupe des MEE.

7.4 Perspectives

L'information contenue dans la matrice CANJEM a récemment été mise en application pour assigner des niveaux d'exposition aux pesticides dans une étude américaine sur le cancer de la thyroïde (Zeng et coll., 2017), et aux métaux et métalloïdes dans le cadre de l'étude internationale INTEROCC sur le cancer du cerveau (Pasquet et coll., 2016). Elle a également été utilisée pour évaluer un outil permettant l'attribution de titres d'emploi à partir de descriptions narratives (Russ et coll., 2016). CANJEM figure également parmi les sources d'informations envisagées pour évaluer l'exposition professionnelle dans deux cohortes populationnelles, l'une en France avec un objectif de 200 000 sujets (Goldberg et coll., 2017), et l'autre Québec pour l'étude CARTaGENE regroupant approximativement 20 000 sujets au total (Awadalla et coll., 2013).

Les efforts de modélisation réalisés dans le cadre de cette thèse ont permis de bonifier l'information contenue dans CANJEM sous deux dimensions, soit le raffinement des estimations pour les cellules moins bien documentées, et l'estimation de niveaux quantitatifs d'intensité de l'exposition. Pour la première dimension, le domaine d'application du cadre de modélisation pourrait être élargi afin d'englober l'ensemble des agents représentés dans CANJEM. Les modèles hiérarchiques basés sur la classification CCDP pourraient également être adaptés aux autres classifications professionnelles et industrielles standardisées disponibles dans CANJEM. Les estimations quantitatives des niveaux d'exposition aux poussières de bois pourraient quant à elle servir à raffiner les connaissances et la prévention des risques à la santé connus (p.ex. cancer des sinus) et soupçonnés (p.ex. cancer du poumon) associés à cet agent. L'application de l'approche développée pour estimer des niveaux quantitatifs d'intensité d'exposition à d'autres agents représente par ailleurs un autre développement souhaitable. La disponibilité de CANJEM sous plusieurs systèmes de classification facilite également l'utilisation d'autres grandes banques de données nationales pour développer des estimations quantitatives de l'exposition adaptées à différents pays.

Les comparaisons menées entre les expositions assignées par l'approche par expertise hybride et celles assignées par expertise traditionnelle ont montré l'utilité des expertises passées comme source d'information pour améliorer la fiabilité dans les estimations tout en tenant compte des caractéristiques propres à chaque emploi. La mise en œuvre de l'approche par expertise hybride nécessite toutefois la disponibilité d'une masse critique d'information sur l'exposition servant de référence pour les experts. À ce titre, les estimations de la matrice CANJEM, disponibles gratuitement en ligne, peuvent représenter un point de départ dans l'application de l'approche par expertise hybride par d'autres groupes de recherche, qui peuvent être ensuite modulés ou bonifiés par un apport d'expertise au contexte de la population d'étude. Par ailleurs, les données d'exposition de l'étude montréalaise PROtEuS représentent une source d'information potentielle pour bonifier l'exposition actuellement contenue dans CANJEM sous deux dimensions. D'une part, elles pourraient permettre de mieux caractériser l'exposition pour les périodes plus récentes. D'autre part, ceci permettrait de fournir des estimations pour les professions absentes de CANJEM, ou d'augmenter la taille d'échantillon pour les autres professions, en particulier pour les professions de cols blancs qui formaient près de la moitié des emplois représentés dans l'étude PROtEuS.

7.5 Conclusion générale

L'exploitation des expertises issues des études cas-témoins montréalaises a permis dans un premier temps de décrire une méthode permettant la synthèse des informations originalement sous une forme individuelle pour former une MEE multidimensionnelle représentant une source d'information unique sur l'exposition rétrospective en milieu de travail pour plus de 250 agents. Dans un contexte de rareté de sources d'information sur l'exposition passée à de multiples agents et couvrant un large spectre de professions et de secteurs industriels, en particulier en Amérique du Nord, CANJEM représente ainsi une référence pouvant être utilisée dans une diversité d'applications en santé au travail et en épidémiologie. Les travaux réalisés ont en outre visé à bonifier les informations en augmentant la précision des estimations et en développant une dimension quantitative. Finalement, la comparaison des expositions assignées avec l'approche par expertise hybride a permis d'observer une augmentation de la fiabilité dans les évaluations comparativement à la méthode par expertise traditionnelle, au prix d'une certaine diminution de la variabilité intra-profession.

Les travaux réalisés s'inscrivent au confluent de deux tendances fortes dans le domaine de l'évaluation de l'exposition dans les études de population. D'abord, il faut viser à une amélioration de l'efficacité de l'évaluation tout en conservant autant que possible les caractéristiques uniques à chaque emploi dans les estimations. De plus, il est souhaitable d'appliquer davantage des méthodes quantitatives afin de raffiner les associations entre l'exposition et les maladies d'origine professionnelle et ainsi mener à une meilleure caractérisation des risques à la santé et à leur prévention.

Bibliographie

Armstrong, B. (2003) «Exposure measurement error: consequences and design issues», dans M.J. Nieuwenhuijsen (sous la direction de), *Exposure Assessment in Occupational and Environmental Epidemiology*, Oxford, Oxford University Press, p. 181-200

Attfield, M.D., P.L. Schleiff, J.H. Lubin, A. Blair, P.A. Stewart, R. Vermeulen, J.B. Coble et D.T. Silverman (2012). «The Diesel Exhaust in Miners study: a cohort mortality study with emphasis on lung cancer», *Journal of the National Cancer Institute*, vol. 104, n° 11, p. 869-883

Awadalla, P., C. Boileau, Y. Payette, Y. Idaghdour, J.-P. Goulet, B. Knoppers, P. Hamet et C. Laberge (2013). «Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics», *International Journal of Epidemiology*, vol. 42, n° 5, p. 1285-1299

Axelson, O. (1979). «The case-referent (case-control) study in occupational health epidemiology», *Scandinavian Journal of Work, Environment and Health*, vol. 5, n° 2, p. 91-99

Baumgarten, M., J. Siemiatycki et G.W. Gibbs (1983). «Validity of Work Histories Obtained By Interview for Epidemiologic Purposes», *American Journal of Epidemiology*, vol. 118, n° 4, p. 583-591

Behrens, T., B. Mester et L. Fritschi (2012). «Sharing the knowledge gained from occupational cohort studies: a call for action», *Occupational and Environmental Medicine*, vol. 69, n° 6, p. 444-448

Bencko, V. (2011). «Use of epidemiological data for cancer risk assessment: approaches, concepts, and issues», *Open Epidemiology Journal*, vol. 4, p. 94-98

Benke, G., M. Sim, L. Fritschi, G. Aldred, A. Forbes et T. Kauppinen (2001). «Comparison of Occupational Exposure Using Three Different Methods: Hygiene Panel, Job Exposure Matrix (JEM), and Self Reports», *Applied Occupational and Environmental Hygiene*, vol. 16, n° 1, p. 84-91

Berglund, M., C.-G. Elinder et L. Järup (2001). *Human exposure assessment: an introduction*, Genève, World Health Organization

Bhatti, P., P.A. Stewart, M.S. Linet, A. Blair, P.D. Inskip et P. Rajaraman (2011). «Comparison of occupational exposure assessment methods in a case–control study of lead, genetic susceptibility and risk of adult brain tumours», *Occupational and Environmental Medicine*, vol. 68, n° 1, p. 4-9

Black, J., G. Benke, K. Smith et L. Fritschi (2004). «Artificial Neural Networks and Job-specific Modules to Assess Occupational Exposure», *Annals of Occupational Hygiene*, vol. 48, n° 7, p. 595-600

Blair, A., R.B. Hayes, P.A. Stewart et S.H. Zahm (1996). «Occupational epidemiologic study design and application», *Occupational Medicine*, vol. 11, n° 3, p. 403-419

Blanc-Lapierre, A., A. Spence, P.I. Karakiewicz, A. Aprikian, F. Saad et M.-É. Parent (2015). «Metabolic syndrome and prostate cancer risk in a population-based case–control study in Montreal, Canada», *BMC Public Health*, vol. 15, n° 1, p. 913

Bosch de Basea, M., M. Porta, J. Alguacil, E. Puigdomenech, M. Gasull, J.A. Garrido et T. Lopez (2011). «Relationships between occupational history and serum concentrations of organochlorine compounds in exocrine pancreatic cancer», *Occupational and Environmental Medicine*, vol. 68, n° 5, p. 332-338

Bourbonnais, R., F. Meyer et G. Thériault (1988). «Validity of self reported work history», *British Journal of Industrial Medicine*, vol. 45, n° 1, p. 29-32

Bouyer, J. et D. Hémon (1993). «Retrospective Evaluation of Occupational Exposures in Population-Based Case-Control Studies: General Overview with Special Attention to Job Exposure Matrices», *International Journal of Epidemiology*, vol. 22, Supplement 2, p. S57-S64

Breslow, N.E. et N.E. Day (sous la direction de) (1987). *Statistical Methods in Cancer Research. Volume II - The Design and Analysis of Cohort Studies* Lyon, International Agency for Research on Cancer

Brodeur, M. et D. Galarneau (1994). «Three large urban areas in transition», *Perspectives on Labour and Income*, vol. 6, n° 4, p. 5

Burns, D.K. et P.L. Beaumont (1989). «The HSE National Exposure Database - (NEDB)», *Annals of Occupational Hygiene*, vol. 33, n° 1, p. 1-14

Burstyn, I., H. Kromhout, P.J. Cruise et P. Brennan (2000). «Designing an International Industrial Hygiene Database of Exposures Among Workers in the Asphalt Industry», *Annals of Occupational Hygiene*, vol. 44, n° 1, p. 57-66

Burstyn, I., J. Lavoué et M. Van Tongeren (2012). «Aggregation of Exposure Level and Probability into a Single Metric in Job-Exposure Matrices Creates Bias», *Annals of Occupational Hygiene*, vol. 56, n° 9, p. 1038-1050

Carey, R.N., T.R. Driscoll, S. Peters, D.C. Glass, A. Reid, G. Benke et L. Fritschi (2014). «Estimated prevalence of exposure to occupational carcinogens in Australia (2011–2012)», *Occupational and Environmental Medicine*, vol. 71, n° 1, p. 55-62

Case, R.A., M.E. Hosker, D.D. Mc et J.T. Pearson (1954). «Tumours of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the British chemical industry. I. The role of aniline, benzidine, alpha-naphthylamine, and beta-naphthylamine», *British Journal of Industrial Medicine*, vol. 11, n° 2, p. 75-104

Centre international de recherche sur le cancer (2006). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans – Preamble*. Lyon, Centre international de Recherche sur le Cancer

Centre international de recherche sur le cancer (2014). *Diesel and Gasoline Engine Exhausts and Some Nitroarenes. Iarc Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 105*, Lyon, Centre international de recherche sur le cancer

Centre international de recherche sur le cancer (2017) Agents Classified by the IARC Monographs, Volumes 1–117. Site Web consulté le 15 mars 2017. http://monographs.iarc.fr/ENG/Classification/

Checkoway, H., N. Pearce et J.M. Dement (1989). «Design and conduct of occupational epidemiology studies: I. design aspects of cohort studies», *American Journal of Industrial Medicine*, vol. 15, n° 4, p. 363-373

Checkoway, H., N. Pearce et D. Kriebel (2004). *Research methods in occupational epidemiology*, Oxford, Oxford University Press

Checkoway, H., N. Pearce et D. Kriebel (2007). «Selecting appropriate study designs to address specific research questions in occupational epidemiology», *Occupational and Environmental Medicine*, vol. 64, n° 9, p. 633-638

Daniels, R.D., T.L. Kubale, J.H. Yiin, M.M. Dahm, T.R. Hales, D. Baris, S.H. Zahm, J.J. Beaumont, K.M. Waters et L.E. Pinkerton (2014). «Mortality and cancer incidence in a pooled cohort of US firefighters from San Francisco, Chicago and Philadelphia (1950–2009)», *Occupational and Environmental Medicine*, vol. 71, n° 6, p. 388

de Vocht, F., J.-P. Zock, H. Kromhout, J. Sunyer, J.M. Antó, P. Burney et M. Kogevinas (2005). «Comparison of self-reported occupational exposure with a job exposure matrix in an international community-based study on asthma», *American Journal of Industrial Medicine*, vol. 47, n° 5, p. 434-442

DellaValle, C.T., M.P. Purdue, M.H. Ward, S.J. Locke, P.A. Stewart, A.J. De Roos, P. Hartge, N. Rothman et M.C. Friesen (2015). «Validity of Expert Assigned Retrospective Estimates of Occupational Polychlorinated Biphenyl Exposure», *The Annals of Occupational Hygiene*, vol. 59, n° 5, p. 609-615

Demers, P., C. Peters, H. Davies, J. Kim, M. Pahwa, C. McLeod, A.-M. Nicol, F. Labrêche, J. Lavoué, S. Hutchings et L. Rushton (2014). «Incorporating more detailed exposure assessment with quantitative estimates is assessing the burden of occupational cancer», *Occupational and Environmental Medicine*, vol. 71, Suppl 1, p. A51-A51 [Abrégé de conférence]

Dewar, R., J. Siemiatycki et M. Gérin (1991). «Loss of Statistical Power Associated with the Use of a Job-Exposure Matrix in Occupational Case-Control Studies», *Applied Occupational and Environmental Hygiene*, vol. 6, n° 6, p. 508-515

Driscoll, T.R., R.N. Carey, S. Peters, D.C. Glass, G. Benke, A. Reid et L. Fritschi (2016). «The Australian Work Exposures Study: Occupational Exposure to Polycyclic Aromatic Hydrocarbons», *Annals of Occupational Hygiene*, vol. 60, n° 1, p. 124-131

Drolet, D. et G. Beauchamp (2012). *Guide d'échantillonnage des contaminants de l'air en milieu de travail (8e édition, version 8.1 mise à jour)*, Guide technique T-06. Montréal, IRSST

EPA (2005). *Guidelines for Carcinogen Risk Assessment*, Washington, U.S. Environmental Protection Agency

Févotte, J., B. Dananché, L. Delabre, S. Ducamp, L. Garras, M. Houot, D. Luce, E. Orlowski, C. Pilorget, A. Lacourt, P. Brochard, M. Goldberg et E. Imbernon (2011). «Matgene: A Program to Develop Job-Exposure Matrices in the General Population in France», *Annals of Occupational Hygiene*, vol. 55, n° 8, p. 865-878

Flegal, K.M., C. Brownie et J. Haas (1986). «The Effects of Exposure Misclassification on Estimates of Relative Risk», *American Journal of Epidemiology*, vol. 123, n° 4, p. 736-751

Friesen, M.C., J.B. Coble, H.A. Katki, B.-T. Ji, S. Xue, W. Lu et P.A. Stewart (2011). «Validity and Reliability of Exposure Assessors' Ratings of Exposure Intensity by Type of Occupational Questionnaire and Type of Rater», *Annals of Occupational Hygiene*, vol. 55, n° 6, p. 601-611

Friesen, M.C., J.B. Coble, W. Lu, X.-O. Shu, B.-T. Ji, S. Xue, L. Portengen, W.-H. Chow, Y.-T. Gao, G. Yang, N. Rothman et R. Vermeulen (2012). «Combining a Job-Exposure Matrix with Exposure Measurements to Assess Occupational Exposure to Benzene in a Population Cohort in Shanghai, China», *Annals of Occupational Hygiene*, vol. 56, n° 1, p. 80-91

Friesen, M.C., S.J. Locke, C. Tornow, Y.-C. Chen, D.-H. Koh, P.A. Stewart, M. Purdue et J.S. Colt (2014). «Systematically Extracting Metal- and Solvent-Related Occupational Information from Free-Text Responses to Lifetime Occupational History Questionnaires», *Annals of Occupational Hygiene*, vol. 58, n° 5, p. 612-624

Friesen, M.C., Y.C. Macnab, S.A. Marion, P.A. Demers, H.W. Davies et K. Teschke (2006). «Mixed Models and Empirical Bayes Estimation for Retrospective Exposure Assessment of Dust Exposures in Canadian Sawmills», *Annals of Occupational Hygiene*, vol. 50, n° 3, p. 281-288

Friesen, M.C., A. Pronk, D.C. Wheeler, Y.-C. Chen, S.J. Locke, D.D. Zaebst, M. Schwenn, A. Johnson, R. Waddell, D. Baris, J.S. Colt, D.T. Silverman, P.A. Stewart et H.A. Katki (2013). «Comparison of Algorithm-based Estimates of Occupational Diesel Exhaust Exposure to Those of Multiple Independent Raters in a Population-based Case–Control Study», *Annals of Occupational Hygiene*, vol. 57, n° 4, p. 470-481

Fritschi, L., M.C. Friesen, D. Glass, G. Benke, J. Girschik et T. Sadkowsky (2009). «OccIDEAS: Retrospective Occupational Exposure Assessment in Community-Based Studies Made Easier», *Journal of Environmental and Public Health*, vol. 2009, Article ID: 957023,

Fritschi, L., L. Nadon, G. Benke, R. Lakhani, B. Latreille, M.-E. Parent et J. Siemiatycki (2003). «Validation of expert assessment of occupational exposures», *American Journal of Industrial Medicine*, vol. 43, n° 5, p. 519-522

Fritschi, L., T. Sadkowsky, G.P. Benke, A. Thomson et D.C. Glass (2012). «Triaging Jobs in a Community-Based Case–Control Study to Increase Efficiency of the Expert Occupational Assessment Method», *Annals of Occupational Hygiene*, vol. 56, n° 4, p. 458-465

Froines, J., D. Wegman et E. Eisen (1989). «Hazard Surveillance in Occupational Disease», *American Journal of Public Health*, vol. 79, Suppl, p. 26-31

Gabriel, S., D. Koppisch et D. Range (2010). «The MGU – a monitoring system for the collection and documentation of valid workplace exposure data », *Gefahrstoffe - Reinhaltung der Luft*, vol. 70, n° 1/2, p. 43-49

Galarneau, D. (2010). «Temporary employment in the downturn», *Perspectives on Labour and Income*, Novembre 2010, p. 5-17

García, A.M. et H. Checkoway (2003). «A glossary for research in occupational health», *Journal of Epidemiology and Community Health*, vol. 57, n° 1, p. 7-10

García, A.M., M.C. González-Galarzo, T. Kauppinen, G.L. Delclos et F.G. Benavides (2013). «A job-exposure matrix for research and surveillance of occupational health and safety in Spanish workers: MatEmESp», *American Journal of Industrial Medicine*, vol. 56, n° 10, p. 1226-1238

Ge, B., C. Peters et P. Demers (2013). «Occupational exposures to known and suspected carcinogens in the Canadian construction industry», *Occupational and Environmental Medicine*, vol. 70, Suppl. 1, p. A109-A110 [Abrégé de conférence]

Gérin, M. et J. Siemiatycki (1991). «The Occupational Questionnaire in Retrospective Epidemiologic Studies: Recent Approaches in Community-Based Studies», *Applied Occupational and Environmental Hygiene*, vol. 6, n° 6, p. 495-499

Gérin, M., J. Siemiatycki, H. Kemper et D. Bégin (1985). «Obtaining occupational exposure histories in epidemiologic case-control studies», *Journal of Occupational Medicine*, vol. 27, n° 6, p. 420-426

Giuffrida, A., R.F. Iunes et W.D. Savedoff (2002). «Occupational risks in Latin America and the Caribbean: economic and health dimensions», *Health Policy and Planning*, vol. 17, n° 3, p. 235-246

Goldberg, M., M. Carton, A. Descatha, A. Leclerc, Y. Roquelaure, G. Santin et M. Zins (2017). «CONSTANCES: a general prospective population-based cohort for occupational and environmental epidemiology: cohort profile», *Occupational and Environmental Medicine*, Prépublication en ligne, 10.1136/oemed-2016-103678

Goldberg, M. et D. Hémon (1993). «Occupational Epidemiology and Assessment of Exposure», *International Journal of Epidemiology*, vol. 22, Supplement 2, p. S5-S9

Goldberg, M. et E. Imbernon (2002). «The use of job exposure matrices for cancer epidemiology research and surveillance», *Archives of Public Health*, vol. 60, n° 3-4, p. 173-185

Goldberg, M., H. Kromhout, P. Guénel, A.C. Fletcher, M. Gérin, D.C. Glass, D. Heederik, T. Kauppinen et A. Ponti (1993). «Job Exposure Matrices in Industry», *International Journal of Epidemiology*, vol. 22, Supplement 2, p. S10-S15

Goldberg, M.S., J. Siemiatycki et M. Gérin (1986). «Inter-rater agreement in assessing occupational exposure in a case-control study», *British Journal of Industrial Medicine*, vol. 43, n° 10, p. 667-676

Gomez, M.R., P. Cocco, M. Dosemeci et P.A. Stewart (1994). «Occupational exposure to chlorinated aliphatic hydrocarbons: Job exposure matrix», *American Journal of Industrial Medicine*, vol. 26, n° 2, p. 171-183

Greenland, S. (2000). «Principles of multilevel modelling», *International Journal of Epidemiology*, vol. 29, n° 1, p. 158-167

Hall, A., P. Demers, C. Peters et H. Davies (2011). «Establishment of a national Canadian workplace exposure database: progress and challenges», *Occupational and Environmental Medicine*, vol. 68, Suppl 1, p. A26

Hall, D.B. (2000). «Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study», *Biometrics*, vol. 56, n° 4, p. 1030-1039

Hardt, J.S., R. Vermeulen, S. Peters, H. Kromhout, J.R. McLaughlin et P.A. Demers (2014). «A comparison of exposure assessment approaches: lung cancer and occupational asbestos exposure in a population-based case–control study», *Occupational and Environmental Medicine*, vol. 71, n° 4, p. 282-288

Hein, M.J., M.A. Waters, A.M. Ruder, M.R. Stenzel, A. Blair et P.A. Stewart (2010). «Statistical modeling of occupational chlorinated solvent exposures for case-control studies using a literature-based database», *Annals of Occupational Hygiene*, vol. 54, n° 4, p. 459-472

Helsel, D. (2010). «Much Ado About Next to Nothing: Incorporating Nondetects in Science», *Annals of Occupational Hygiene*, vol. 54, n° 3, p. 257-262

Hewett, P., P. Logan, J. Mulhausen, G. Ramachandran et S. Banerjee (2006). «Rating Exposure Control Using Bayesian Decision Analysis», *Journal of Occupational and Environmental Hygiene*, vol. 3, n° 10, p. 568-581

Hoar, S. (1983). «Job Exposure Matrix Methodology», *Clinical Toxicology*, vol. 21, n° 1-2, p. 9-26

Hoar, S.K., A.S. Morrison, P. Cole et D.T. Silverman (1980). «An occupation and exposure linkage system for the study of occupational carcinogenesis», *Journal of Occupational Medicine*, vol. 22, n° 11, p. 722-726

Hobson, J. (2013). «To die for? The health and safety of fast fashion», *Occupational Medicine*, vol. 63, n° 5, p. 317-319

Hogstedt, C. et B. Pieris (2000). *Occupational Safety and Health in Developing Countries. Review of strategies, case studies and a bibliography.* Arbetslivsrapport 2000:17, Stochholm, Arbeidslivsinstituttet

Hon, C.-Y., C.E. Peters, K.J. Jardine et V.H. Arrandale (2017). «Historical occupational isocyanate exposure levels in two Canadian provinces», *Journal of Occupational and Environmental Hygiene*, vol. 14, n° 1, p. 1-8

Huynh, T., H. Quick, G. Ramachandran, S. Banerjee, M. Stenzel, D.P. Sandler, L.S. Engel, R.K. Kwok, A. Blair et P.A. Stewart (2016). «A Comparison of the β-Substitution Method and a Bayesian Method for Analyzing Left-Censored Data», *The Annals of Occupational Hygiene*, vol. 60, n° 1, p. 56-73

Kachuri, L., P.J. Villeneuve, M.-É. Parent, K.C. Johnson et S.A. Harris (2016). «Workplace exposure to diesel and gasoline engine exhausts and the risk of colorectal cancer in Canadian men», *Environmental Health*, vol. 15, n° 1, p. 4

Kauppinen, T. (1996). «Exposure assessment – A challenge for occupational epidemiology», *Scandinavian Journal of Work, Environment and Health*, vol. 22, n° 6, p. 401-403

Kauppinen, T., P. Heikkilä, N. Plato, T. Woldbæk, k. lenvik, J. Hansen, V. Kristjansson et E. Pukkala (2009). «Construction of job-exposure matrices for the Nordic Occupational Cancer Study (NOCCA)», *Acta Oncologica*, vol. 48, n° 5, p. 791-800

Kauppinen, T., J. Toikkanen et E. Pukkala (1998). «From cross-tabulations to multipurpose exposure information systems: A new job-exposure matrix», *American Journal of Industrial Medicine*, vol. 33, n° 4, p. 409-417

Kauppinen, T., S. Uuksulainen, A. Saalo, I. Mäkinen et E. Pukkala (2014). «Use of the Finnish Information System on Occupational Exposure (FINJEM) in Epidemiologic, Surveillance, and Other Applications», *Annals of Occupational Hygiene*, vol. 58, n° 3, p. 380-396

Kielhorn, J., C. Melber, U. Wahnschaffe, A. Aitio et I. Mangelsdorf (2000). «Vinyl chloride: still a cause for concern», *Environmental Health Perspectives*, vol. 108, n° 7, p. 579-588

Kirkham, T.L., J. Siemiatycki, F. Labrèche et J. Lavoué (2016). «Impact of aggregating exposure information from cases and controls when building a population-based job-exposure matrix from past expert evaluations», *Occupational and Environmental Medicine*, vol. 73, n° 7, p. 474-481

Koeman, T., N. S. M. Offermans, Y. Christopher-de Vries, P. Slottje, P.A. Van Den Brandt, R.A. Goldbohm, H. Kromhout et R. Vermeulen (2013). «JEMs and Incompatible Occupational Coding Systems: Effect of Manual and Automatic Recoding of Job Codes on Exposure Assignment», *Annals of Occupational Hygiene*, vol. 57, n° 1, p. 107-114

Koh, D.-H., P. Bhatti, J.B. Coble, P.A. Stewart, W. Lu, X.-O. Shu, B.-T. Ji, S. Xue, S.J. Locke, L. Portengen, G. Yang, W.-H. Chow, Y.-T. Gao, N. Rothman, R. Vermeulen et M.C. Friesen (2014). «Calibrating a population-based job-exposure matrix using inspection measurements to estimate historical occupational exposure to lead for a population-based cohort in Shanghai, China», *Journal of Exposure Science and Environmental Epidemiology*, vol. 24, n° 1, p. 9-16

Kromhout, H. (2002). «Commentary», *Occupational and Environmental Medicine*, vol. 59, n° 9, p. 594

Kromhout, H., Y. Oostendorp, D. Heederik et J.S.M. Boleij (1987). «Agreement between qualitative exposure estimates and quantitative exposure measurements», *American Journal of Industrial Medicine*, vol. 12, n° 5, p. 551-562

Kromhout, H. et R. Vermeulen (2001). «Application of job-exposure matrices in studies of the general population-some clues to their performance», *European Respiratory Review*, vol. 11, n° 80, p. 80-90

Kumagai, S., N. Kurumatani, A. Arimoto et G. Ichihara (2013). «Cholangiocarcinoma among offset colour proof-printing workers exposed to 1,2-dichloropropane and/or dichloromethane», *Occupational and Environmental Medicine*, vol. 70, n° 7, p. 508-510

Labrèche, F., P. Duguay, C. Ostiguy, N. Goyer, A. Boucher, B. Roberge et M. Baril (2012). *Substances cancérogènes – Portrait de l'exposition des travailleurs québécois*, Études et recherches, Rapport R-732. Montréal, Institut de recherche Robert-Sauvé en santé et en sécurité du travail

Labrèche, F., M.S. Goldberg, M.-F. Valois et L. Nadon (2010). «Postmenopausal breast cancer and occupational exposures», *Occupational and Environmental Medicine*, vol. 67, n° 4, p. 263-269

Labrèche, F., A. Lacourt et J. Lavoué (2014). *Expositions professionnelles à des contaminants chimiques et physiques: Analyse différenciée selon le sexe*, Études et recherches, Rapport R-842. Montréal, Institut de recherche Robert-Sauvé en santé et en sécurité du travail

Lacourt, A., E. Cardis, J. Pintos, L. Richardson, L. Kincl, G. Benke, S. Fleming, M. Hours, D. Krewski, D. McLean, M.-E. Parent, S. Sadetzki, K. Schlaefer, B. Schlehofer, M. van Tongeren et J. Siemiatycki (2013). «INTEROCC case--control study: lack of association between glioma tumors and occupational exposure to selected combustion products, dusts and other chemical agents», *BMC Public Health*, vol. 13, p. 340

Lavoué, J. (2006). *Évaluation de l'exposition professionnelle au formaldéhyde à partir de sources de données préexistantes*, Thèse présentée à la Faculté des études supérieures en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.) en Santé publique, option Toxicologie de l'environnement. Montréal, Université de Montréal

Lavoué, J., D. Bégin, C. Beaudry et M. Gérin (2007). «Monte Carlo Simulation to Reconstruct Formaldehyde Exposure Levels from Summary Parameters Reported in the Literature», *Annals of Occupational Hygiene*, vol. 51, n° 2, p. 161-172

Lavoué, J., M.C. Friesen et I. Burstyn (2012a). «Workplace Measurements by the US Occupational Safety and Health Administration since 1979: Descriptive Analysis and Potential Uses for Exposure Assessment», *Annals of Occupational Hygiene*, vol. 57, n° 1, p. 77-97

Lavoué, J., M. Gérin, D. Bégin, C. Ostiguy, R. Arcand et G. Adib (2012b). *Valorisation des données d'exposition professionnelle mesurées au Québec depuis 1980 par les équipes du Réseau public québécois en santé au travail – Étude préliminaire*, Études et recherches, Rapport R-723. Montréal, Institut de recherche Robert-Sauvé en santé et en sécurité du travail

Lavoué, J., J. Pintos, M. Van Tongeren, L. Kincl, L. Richardson, T. Kauppinen, E. Cardis et J. Siemiatycki (2012c). «Comparison of exposure estimates in the Finnish job-exposure matrix FINJEM with a JEM derived from expert assessments performed in Montreal», *Occupational and Environmental Medicine*, vol. 69, n° 7, p. 465-471

Li, C.Y. et F.C. Sung (1999). «A review of the healthy worker effect in occupational epidemiology», *Occupational Medicine (Oxford, England)*, vol. 49, n° 4, p. 225-229

Lin, Y.S., L.L. Kupper et S.M. Rappaport (2005). «Air samples versus biomarkers for epidemiology», *Occupational and Environmental Medicine*, vol. 62, n° 11, p. 750-760

Luce, D. et J. Févotte (2006). *Le programme Matgéné: Matrices emplois-expositions en population générale*, Saint-Maurice, Institut de Veille Sanitaire

Luce, D., I. Stucker et ICARE study group (2011). «Investigation of occupational and environmental causes of respiratory cancers (ICARE): a multicenter, population-based case-control study in France», *BMC Public Health*, vol. 11, n° 1, p. 928

Lunn, D.J., A. Thomas, N. Best et D. Spiegelhalter (2000). «WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility», *Statistics and computing*, vol. 10, n° 4, p. 325-337

Macfarlane, E., G. Benke, M.R. Sim et L. Fritschi (2012). «OccIDEAS: An Innovative Tool to Assess Past Asbestos Exposure in the Australian Mesothelioma Registry», *Safety and Health at Work*, vol. 3, n° 1, p. 71-76

McGuire, V., L.M. Nelson, T.D. Koepsell, H. Checkoway et W.T. Longstreth (1998). «Assessment of Occupational Exposures in Community-Based Case-Control Studies», *Annual Review of Public Health*, vol. 19, n° 1, p. 35-53

McLean, D., S. Fleming, M.C. Turner, L. Kincl, L. Richardson, G. Benke, B. Schlehofer, K. Schlaefer, M.-E. Parent, M. Hours, D. Krewski, M. van Tongeren, S. Sadetzki, J. Siemiatycki et E. Cardis (2014). «Occupational solvent exposure and risk of meningioma: results from the INTEROCC multicentre case–control study», *Occupational and Environmental Medicine*, vol. 71, n° 4, p. 253-258

Ministères des finances du Québec (2005). «Le vêtement et le textile au Québec: Deux industries face à des défis considérables», *Études économiques, fiscales et budgétaires*, vol. 1, n° 4,

Momoli, F., M. Abrahamowicz, M.E. Parent, D. Krewski et J. Siemiatycki (2010). «Analysis of multiple exposures: an empirical comparison of results from conventional and semi-bayes modeling strategies», *Epidemiology*, vol. 21, n° 1, p. 144-151

Mrema, E.J., A.V. Ngowi et S.H. Mamuya (2015). «Status of Occupational Health and Safety and Related Challenges in Expanding Economy of Tanzania», *Annals of Global Health*, vol. 81, n° 4, p. 538-547

National Research Council (1983). *Risk Assessment in the Federal Government: Managing the Process*, Washington, The National Academies Press

Neilson, H.K., A. Sass-Kortsak, W.Y.W. Lou et J.T. Purdham (2007). «Personal factors influencing agreement between expert and self-reported assessments of an occupational exposure», *Chronic Diseases in Canada*, vol. 28, n° 1-2, p. 1

Nieuwenhuijsen, M., D. Paustenbach et R. Duarte-Davidson (2006). «New developments in exposure assessment: The impact on the practice of health risk assessment and epidemiological studies», *Environment International*, vol. 32, n° 8, p. 996-1009

Nieuwenhuijsen, M.J. (2003) «Introduction to exposure assessment», dans M.J. Nieuwenhuijsen (sous la direction de), *Exposure assessment in occupational and environmental epidemiology*, Oxford, Oxford University Press, p. 3-20

Nieuwenhuijsen, M.J. et P. Droz (2003) «Biological monitoring», dans M.J. Nieuwenhuijsen (sous la direction de), *Exposure assessment in occupational and environmental epidemiology*, Oxford, Oxford University Press, p. 167-180

Nkosi, T.M., M.E. Parent, J. Siemiatycki et M.C. Rousseau (2012). «Socioeconomic position and lung cancer risk: how important is the modeling of smoking?», *Epidemiology*, vol. 23, n° 3, p. 377-385

Offermans, N.S.M., R. Vermeulen, A. Burdorf, S. Peters, R.A. Goldbohm, T. Koeman, M. van Tongeren, T. Kauppinen, I. Kant, H. Kromhout et P.A. van den Brandt (2012). «Comparison of expert and job-exposure matrix-based retrospective exposure assessment of occupational carcinogens in the Netherlands Cohort Study», *Occupational and Environmental Medicine*, vol. 69, n° 10, p. 745-751

Olsen, J. (1988). «Limitations in the use of Job Exposure Matrices», *Scandinavian Journal of Public Health*, vol. 16, n° 4, p. 205-208

Olsson, A.C., P. Gustavsson, H. Kromhout, S. Peters, R. Vermeulen, I. Brüske, B. Pesch, J. Siemiatycki, J. Pintos, T. Brüning, A. Cassidy, H.-E. Wichmann, D. Consonni, M.T. Landi, N. Caporaso, N. Plato, F. Merletti, D. Mirabelli, L. Richiardi, K.-H. Jöckel, W. Ahrens, H. Pohlabeln, J. Lissowska, N. Szeszenia-Dabrowska, D. Zaridze, I. Stücker, S. Benhamou, V. Bencko, L. Foretova, V. Janout, P. Rudnai, E. Fabianova, R.S. Dumitru, I.M. Gross, B. Kendzia, F. Forastiere, B. Bueno-de-Mesquita, P. Brennan, P. Boffetta et K. Straif (2011). «Exposure to Diesel Motor Exhaust and Lung Cancer Risk in a Pooled Analysis from Case-Control Studies in Europe and Canada», *American Journal of Respiratory and Critical Care Medicine*, vol. 183, n° 7, p. 941-948

Pannett, B., D. Coggon et E.D. Acheson (1985). «A job-exposure matrix for use in population based studies in England and Wales», *British Journal of Industrial Medicine*, vol. 42, n° 11, p. 777-783

Parmet, A.J. et S. Von Essen (2002). «Rapidly progressive, fixed airway obstructive disease in popcorn workers: a new occupational pulmonary illness?», *Journal of Occupational and Environmental Medicine*, vol. 44, n° 3, p. 216-218

Pasquet, R., E. Cardis, L. Richardson, J. Lavoué, J. Siemiatycki et A. Koushik (2016). «The association between occupational exposure to metals and metalloids and brain cancer risk», *Occupational and Environmental Medicine*, vol. 73, Suppl 1, p. A134-A135 [Abrégé de conférence]

Paustenbach, D.J. (2000). «The practice of exposure assessment: A state-of-the-art review», *Journal of Toxicology and Environmental Health, Part B*, vol. 3, n° 3, p. 179-291

Pearce, N., H. Checkoway et J. Dement (1989). «Design and conduct of occupational epidemiology studies: III. Design aspects of case-control studies», *American Journal of Industrial Medicine*, vol. 15, n° 4, p. 395-402

Peters, C.E., C.B. Ge, A.L. Hall, H.W. Davies et P.A. Demers (2015). «CAREX Canada: an enhanced model for assessing occupational carcinogen exposure», *Occupational and Environmental Medicine*, vol. 72, n° 1, p. 64-71

Peters, S., R. Vermeulen, A. Cassidy, A.t. Mannetje, M. van Tongeren, P. Boffetta, K. Straif et H. Kromhout (2011a). «Comparison of exposure assessment methods for occupational carcinogens in a multi-centre lung cancer case-control study», *Occupational and Environmental Medicine*, vol. 68, n° 2, p. 148-153

Peters, S., R. Vermeulen, L. Portengen, A. Olsson, B. Kendzia, R. Vincent, B. Savary, J. Lavoué, D. Cavallo, A. Cattaneo, D. Mirabelli, N. Plato, J. Fevotte, B. Pesch, T. Bruning, K. Straif et H. Kromhout (2011b). «Modelling of occupational respirable crystalline silica exposure for quantitative exposure assessment in community-based case-control studies», *Journal of Environmental Monitoring*, vol. 13, n° 11, p. 3262-3268

Peters, S., R. Vermeulen, L. Portengen, A. Olsson, B. Kendzia, R. Vincent, B. Savary, J. Lavoué, D. Cavallo, A. Cattaneo, D. Mirabelli, N. Plato, J. Fevotte, B. Pesch, T. Brüning, K. Straif et H. Kromhout (2016). «SYN-JEM: A Quantitative Job-Exposure Matrix for Five Lung Carcinogens», *Annals of Occupational Hygiene*, vol. 60, n° 7, p. 795-811

Pingle, S. (2012). «Occupational safety and health in India: now and the future», *Industrial Health*, vol. 50, n° 3, p. 167-171

Plato, N. et G. Steineck (1993). «Methodology and utility of a job-exposure matrix», *American Journal of Industrial Medicine*, vol. 23, n° 3, p. 491-502

Pleil, J.D., B.C. Blount, S. Waidyanatha et M. Harper (2012). «Establishing exposure science as a distinct scientific discipline», *Journal of Exposure Science and Environmental Epidemiology*, vol. 22, n° 4, p. 317-319

Plummer, M. (2003) *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*. In: Proceedings of the 3rd international workshop on distributed statistical computing, pp. 125. Vienne

Pold, H. (2001). «Trends in part-time work», *Perspectives on Labour and Income*, vol. 13, n° 1, p. 36-38

Pronk, A., P.A. Stewart, J.B. Coble, H.A. Katki, D.C. Wheeler, J.S. Colt, D. Baris, M. Schwenn, M.R. Karagas, A. Johnson, R. Waddell, C. Verrill, S. Cherala, D.T. Silverman et M.C. Friesen (2012). «Comparison of two expert-based assessments of diesel exhaust exposure in a case–control study: programmable decision rules versus expert review of individual jobs», *Occupational and Environmental Medicine*, vol. 69, n° 10, p. 752-758

Pukkala, E., J. Guo, P. Kyyronen, M.L. Lindbohm, M. Sallmen et T. Kauppinen (2005). «National job-exposure matrix in analyses of census-based estimates of occupational cancer risk», *Scandinavian Journal of Work, Environment and Health*, vol. 31, n° 2, p. 97-107

Purdue, M.P., P.A. Stewart, M.C. Friesen, J.S. Colt, S.J. Locke, M.J. Hein, M.A. Waters, B.I. Graubard, F. Davis, J. Ruterbusch, K. Schwartz, W.-H. Chow, N. Rothman et J.N. Hofmann (2017). «Occupational exposure to chlorinated solvents and kidney cancer: a case–control study», *Occupational and Environmental Medicine*, vol. 74, n° 4, p. 268-274

Quinn, M.M. (2011). «Why do women and men have different occupational exposures?», *Occupational and Environmental Medicine*, vol. 68, n° 12, p. 861-862

Ramachandran, G. (2001). «Retrospective exposure assessment using bayesian methods», *Annals of Occupational Hygiene*, vol. 45, n° 8, p. 651-667

Ramachandran, G. (2008). «Toward Better Exposure Assessment Strategies--The New NIOSH Initiative», *Annals of Occupational Hygiene*, vol. 52, n° 5, p. 297-301

Ramachandran, G. et J.H. Vincent (1999). «A Bayesian Approach to Retrospective Exposure Assessment», *Applied Occupational and Environmental Hygiene*, vol. 14, n° 8, p. 547-557

Rappaport, S.M. (2009). «Assessing workplace exposures: turning to the past for guidance», *Occupational and Environmental Medicine*, vol. 66, n° 7, p. 429-430

Rappaport, S.M., E. Symanski, J.W. Yager et L.L. Kupper (1995). «The relationship between environmental monitoring and biological markers in exposure assessment», *Environmental Health Perspectives*, vol. 103 Suppl 3, p. 49-53

Rocheleau, C.M., C.C. Lawson, M.A. Waters, M.J. Hein, P.A. Stewart, A. Correa, D. Echeverria et J. Reefhuis (2011). «Inter-Rater Reliability of Assessed Prenatal Maternal Occupational Exposures to Solvents, Polycyclic Aromatic Hydrocarbons, and Heavy Metals», *Journal of Occupational and Environmental Hygiene*, vol. 8, n° 12, p. 718-728

Rothman, K.J. et S. Greenland (2008) «Cohort Studies», dans K.J. Rothman, S. Greenland et T.L. Lash (sous la direction de), *Modern epidemiology*, Philadelphie, Wolters Kluwer | Lippincott Williams & Wilkins, p. 100-111

Russ, D., T. Remen, K.-Y. Ho, W.-H. Chow, F. Davis, J. Hofmann, H. Huang, M. Purdue, K. Schwartz, J. Siemiatycki, Y. Zhang, D. Silverman, C. Johnson, J. Lavoué et M. Friesen (2016). « Recommendations for prioritising expert review of free-text job descriptions that underwent computer-based coding using the soccer algorithm», *Occupational and Environmental Medicine*, vol. 73, Suppl 1, p. A186 [Abrégé de conférence]

Sadhra, S.S., O.P. Kurmi, H. Chambers, K.B.H. Lam et D. Fishwick (2016). «Development of an occupational airborne chemical exposure matrix», *Occupational Medicine*, vol. 66, n° 5, p. 358-364

Sahasranaman, A. (2000). *Occupational Safety and Health in the Tanning Industry in South East Asia*, Fourteenth Session of the Leather and Leather Products Industry Panel. Zlin, Organisation des Nations Unies pour le développement industriel

Sahmel, J., K. Devlin, D. Paustenbach, D. Hollins et S. Gaffney (2010). «The role of exposure reconstruction in occupational human health risk assessment: Current methods and a recommended framework», *Critical Reviews in Toxicology*, vol. 40, n° 9, p. 799-843

Sarazin, P., I. Burstyn, L. Kincl et J. Lavoué (2016). «Trends in OSHA Compliance Monitoring Data 1979–2011: Statistical Modeling of Ancillary Information across 77 Chemicals», *Annals of Occupational Hygiene*, vol. 60, n° 4, p. 432-452

Sauvé, J.-F., J. Lavoué et M.-É. Parent (2016). «Occupation, industry, and the risk of prostate cancer: a case-control study in Montréal, Canada», *Environmental Health*, vol. 15, n° 1, p. 100

Sauvé, J.F., M. Lévesque, M. Huard, D. Drolet, J. Lavoué, R. Tardif et G. Truchon (2015). «Creatinine and specific gravity normalization in biological monitoring of occupational exposures», *Journal of Occupational and Environmental Hygiene*, vol. 12, n° 2, p. 123-129

Scarselli, A., C. Montaruli et A. Marinaccio (2007). «The Italian Information System on Occupational Exposure to Carcinogens (SIREP): Structure, Contents and Future Perspectives», *Annals of Occupational Hygiene*, vol. 51, n° 5, p. 471-478

Semple, S. (2005). «Assessing occupational and environmental exposure», *Occupational Medicine*, vol. 55, n° 6, p. 419-424

Semple, S.E., F. Dick et J.W. Cherrie (2004). «Exposure assessment for a population-based case-control study combining a job-exposure matrix with interview data», *Scandinavian Journal of Work, Environment and Health*, vol. 30, n° 3, p. 241-248

Si, S., R. Carey, A. Reid, S. Peters, D.D. Glass, T. Driscoll, E. Darcey et L. Fritschi (2016). «Occupational exposure to carcinogens in Australian road transport workers», *American Journal of Industrial Medicine*, vol. 59, n° 1, p. 31-41

Siemiatycki, J. (1991). *Risk factors for cancer in the workplace*, Boca Raton, CRC Press

Siemiatycki, J. (1996). «Exposure assessment in community-based studies of occupational cancer», *Occupational Hygiene*, vol. 3, n° 1, p. 41-58

Siemiatycki, J. (2007). «Investigating cancer risks related to asbestos and other occupational carcinogens», *Occupational and Environmental Medicine*, vol. 64, n° 8, p. 500-501

Siemiatycki, J., N.E. Day, J. Fabry et J.A. Cooper (1981). «Discovering Carcinogens in the Occupational Environment: A Novel Epidemiologic Approach», *Journal of the National Cancer Institute*, vol. 66, n° 2, p. 217-225

Siemiatycki, J., R. Dewar et L. Richardson (1989). «Costs and Statistical Power Associated With Five Methods of Collecting Occupation Exposure Information for Population-Based Case-Control Studies», *American Journal of Epidemiology*, vol. 130, n° 6, p. 1236-1246

Siemiatycki, J., L. Fritschi, L. Nadon et M. Gérin (1997). «Reliability of an expert rating procedure for retrospective assessment of occupational exposures in community-based case-control studies», *American Journal of Industrial Medicine*, vol. 31, n° 3, p. 280-286

Siemiatycki, J., L. Nadon, R. Lakhani, D. Bégin et M. Gérin (1991) «Exposure assessment», dans J. Siemiatycki (sous la direction de), *Risk factors for cancer in the workplace*, Boca Raton, CRC Press, p. 45-114

Siemiatycki, J. et L. Richardson (1991) «Case-control design and fieldwork methods», dans J. Siemiatycki (sous la direction de), *Risk factors for cancer in the workplace*, Boca Raton, CRC Press, p. 29-44

Siemiatycki, J., L. Richardson et P. Boffetta (2006) «Occupation», dans D. Schottenfeld et J.F. Fraumeni (sous la direction de), *Cancer epidemiology and prevention*, Oxford, Oxford University Press, p. 322-354

Siemiatycki, J., L. Richardson, K. Straif, B. Latreille, R. Lakhani, S. Campbell, M.-C. Rousseau et P. Boffetta (2004). «Listing occupational carcinogens», *Environmental Health Perspectives*, vol. 112, n° 15, p. 1447-1459

Siemiatycki, J., S. Wacholder, L. Richardson, R. Dewar et M. Gérin (1987). «Discovering carcinogens in the occupational environment. Methods of data collection and analysis of a large case-referent monitoring system», *Scandinavian Journal of Work, Environment and Health*, vol. 13, n° 6, p. 486-492

Smith, T.J., P.A. Stewart et R.F. Herrick (2005) «Retrospective exposure assessment», dans K. Gardiner et J.M. Harrington (sous la direction de), *Occupational Hygiene, 3rd Edition*, Oxford, Blackwell Publishing, p. 145-159

Statistics Canada (2003). *North American Industry Classification System, Canada 2002*, Ottawa, Statistics Canada

Statistics Canada (2007). *National Occupational Classification for Statistics (NOC-S)*, Ottawa, Statistics Canada

Steenland, K., I. Bray, S. Greenland et P. Boffetta (2000). «Empirical Bayes Adjustments for Multiple Results in Hypothesis-generating or Surveillance Studies», *Cancer Epidemiology Biomarkers & Prevention*, vol. 9, n° 9, p. 895-903

Steenland, K. et J.A. Deddens (2004). «A Practical Guide to Dose-Response Analyses and Risk Assessment in Occupational Epidemiology», *Epidemiology*, vol. 15, n° 1, p. 63-70

Steenland, K., A. Mannetje, P. Boffetta, L. Stayner, M. Attfield, J. Chen, M. Dosemeci, N. DeKlerk, E. Hnizdo, R. Koskela et H. Checkoway (2001). «Pooled exposure–response analyses and risk assessment for lung cancer in 10 cohorts of silica-exposed workers: an IARC multicentre study», *Cancer Causes and Control*, vol. 12, n° 9, p. 773-784

Stewart, P. (1999). «Challenges to retrospective exposure assessment», *Scandinavian Journal of Work, Environment and Health*, vol. 25, n° 6, p. 505-510

Stewart, P. et M. Stenzel (1999). «Viewpoint . Data needs for occupational epidemiologic studies», *Journal of Environmental Monitoring*, vol. 1, n° 4, p. 75N-82N

Stewart, P.A., P.S. Lees et M. Francis (1996). «Quantification of historical exposures in occupational cohort studies», *Scandinavian Journal of Work, Environment and Health*, vol. 22, n° 6, p. 405-414

Stewart, P.A. et C. Rice (1990). «A Source of Exposure Data for Occupational Epidemiology Studies», *Applied Occupational and Environmental Hygiene*, vol. 5, n° 6, p. 359-363

Stewart, W.F. et P.A. Stewart (1994). «Occupational case-control studies: I. Collecting information on work histories and work-related exposures», *American Journal of Industrial Medicine*, vol. 26, n° 3, p. 297-312

t' Mannetje, A., K. Steenland, H. Checkoway, R.-S. Koskela, M. Koponen, M. Attfield, J. Chen, E. Hnizdo, N. DeKlerk et M. Dosemeci (2002). «Development of quantitative exposure data for a pooled exposure-response analysis of 10 silica cohorts», *American Journal of Industrial Medicine*, vol. 42, n° 2, p. 73-86

t' Mannetje, A.M., D.J. McLean, A.J. Eng, H. Kromhout, T. Kauppinen, J. Fevotte et N.E. Pearce (2011). «Developing a General Population Job-Exposure Matrix in the Absence of Sufficient Exposure Monitoring Data», *Annals of Occupational Hygiene*, vol. 55, n° 8, p. 879-885

ter Burg, W. (2014). *Inventory of databases containing worker exposure data on non-threshold carcinogens in Europe. RIVM Letter report 2014-0083*, Bilthoven, National Institute for Public Health and the Environment

Teschke, K. (2003) «Exposure surrogates: job-exposure matrices, self-reports, and expert evaluations», dans M.J. Nieuwenhuijsen (sous la direction de), *Exposure Assessment in Occupational and Environmental Epidemiology*, Oxford, Oxford University Press, p. 118-132

Teschke, K., A.F. Olshan, J.L. Daniels, A.J. De Roos, C.G. Parks, M. Schulz et T.L. Vaughan (2002). «Occupational exposure assessment in case-control studies: opportunities for improvement», *Occupational and Environmental Medicine*, vol. 59, n° 9, p. 575-594

Tielemans, E., H. Marquart, J. De Cock, M. Groenewold et J. Van Hemmen (2002). «A Proposal for Evaluation of Exposure Data», *Annals of Occupational Hygiene*, vol. 46, n° 3, p. 287-297

Vadali, M., G. Ramachandran et J. Mulhausen (2009). «Exposure Modeling in Occupational Hygiene Decision Making», *Journal of Occupational and Environmental Hygiene*, vol. 6, n° 6, p. 353-362

van Tongeren, M., L. Kincl, L. Richardson, G. Benke, J. Figuerola, T. Kauppinen, R. Lakhani, J. Lavoué, D. McLean, N. Plato et E. Cardis (2013). «Assessing Occupational Exposure to Chemicals in an International Epidemiological Study of Brain Tumours», *Annals of Occupational Hygiene*, vol. 57, n° 5, p. 610-626

Vermeulen, R. (2016). «When Are Risk Analyses on Job Titles Informative?», *Annals of Occupational Hygiene*, vol. 60, n° 8, p. 913-915

Verner, M.A., D. Bachelet, R. McDougall, M. Charbonneau, P. Guenel et S. Haddad (2011). «A case study addressing the reliability of polychlorinated biphenyl levels measured at the time of breast cancer diagnosis in representing early-life exposure», *Cancer Epidemiology Biomarkers & Prevention*, vol. 20, n° 2, p. 281-286

Viet, S.M., M. Stenzel, C.P. Rennix, T.W. Armstrong et J.R. Couch (2008). *Guideline on occupational exposure reconstruction*, Fairfax, AIHA Press

Vincent, R. et B. Jeandel (2001). «COLCHIC - Occupational Exposure to Chemical Agents Database: Current Content and Development Perspectives», *Applied Occupational and Environmental Hygiene*, vol. 16, n° 2, p. 115 - 121

Vincent, R., T. Kauppinen, J. Toikkanen, D. Pedersen, R. Young et M. Kogevinas (1999). «CAREX: Système international d'information sur l'exposition professionnelle aux agents cancérogènes en Europe. Résultats des estimations pour la France pendant les années 1990-1993 », *Cahiers de notes documentaires - Hygiène et sécurité du travail*, n° 176, p. 48-58

Vineis, P. et A. Blair (1992). «Problems and perspectives in the identification of new occupational carcinogens», *Scandinavian Journal of Work, Environment and Health*, vol. 18, n° 5, p. 273-277

Wheeler, D.C., K.J. Archer, I. Burstyn, K. Yu, P.A. Stewart, J.S. Colt, D. Baris, M.R. Karagas, M. Schwenn, A. Johnson, K. Armenti, D.T. Silverman et M.C. Friesen (2015). «Comparison of Ordinal and Nominal Classification Trees to Predict Ordinal Expert-Based Occupational Exposure Estimates in a Case–Control Study», *Annals of Occupational Hygiene*, vol. 59, n° 3, p. 324-335

Wheeler, D.C., I. Burstyn, R. Vermeulen, K. Yu, S.M. Shortreed, A. Pronk, P.A. Stewart, J.S. Colt, D. Baris, M.R. Karagas, M. Schwenn, A. Johnson, D.T. Silverman et M.C. Friesen (2013). «Inside the black box: starting to uncover the underlying decision rules used in a one-by-one expert assessment of occupational exposure in case-control studies», *Occupational and Environmental Medicine*, vol. 70, n° 3, p. 203-210

Wild, P., E.A. Sauleau, E. Bourgkard et J.-J. Moulin (2002). «Combining Expert Ratings and Exposure Measurements: A Random Effect Paradigm», *Annals of Occupational Hygiene*, vol. 46, n° 5, p. 479-487

Zeng, F., C. Lerro, J. Lavoué, H. Huang, J. Siemiatycki, N. Zhao, S. Ma, N.C. Deziel, M.C. Friesen, R. Udelsman et Y. Zhang (2017). «Occupational exposure to pesticides and other biocides and risk of thyroid cancer», *Occupational and Environmental Medicine*, Prépublication en ligne, 10.1136/oemed-2016-103931

Zheng, W., W.-H. Chow, G. Yang, F. Jin, N. Rothman, A. Blair, H.-L. Li, W. Wen, B.-T. Ji, Q. Li, X.-O. Shu et Y.-T. Gao (2005). «The Shanghai Women's Health Study: Rationale, Study Design, and Baseline Characteristics», *American Journal of Epidemiology*, vol. 162, n° 11, p. 1123-1131

Annexe 1. Occupation, industry, and the risk of prostate cancer: a case-control study in Montréal, Canada

Article publié dans *Environmental Health* vol. 15, n° 1, p. 100 (2016)

Jean-François Sauvé [1,2], Jérôme Lavoué [1,2], Marie-Élise Parent [2,3,4]

1. Université de Montréal, School of Public Health, Department of environmental and occupational health, Montréal, Québec, Canada
2. Centre de recherche du CHUM, Montréal, Québec, Canada
3. INRS-Institut Armand-Frappier, Université du Québec, Laval, Québec, Canada
4. Université de Montréal, School of Public Health, Department of social and preventive medicine, Montréal, Québec, Canada

A1.1. Abstract

Background: Age, family history and ancestry are the only recognized risk factors for prostate cancer (PCa) but a role for environmental factors is suspected. Due to the lack of knowledge on the etiological factors for PCa, studies that are both hypothesis-generating and confirmatory are still needed. This study explores relationships between employment, by occupation and industry, and PCa risk.

Methods: Cases were 1937 men aged ≤75 years with incident PCa diagnosed across Montreal French hospitals in 2005-2009. Controls were 1994 men recruited concurrently from electoral lists of French-speaking Montreal residents, frequency-matched to cases by age. In-person interviews elicited occupational histories. Unconditional logistic regression estimated odds ratios (OR) and 95% confidence intervals (CI) for the association between employment across 696 occupations and 613 industries and PCa risk, adjusting for potential confounders. Multinomial logistic models assessed risks by PCa grade. Semi-Bayes (SB) adjustment accounted for the large number of associations evaluated.

Results: Consistently positive associations—and generally robust to SB adjustment—were found for occupations in forestry and logging (OR 1.9, 95%CI: 1.2-3.0), social sciences (OR 1.6, 95%CI: 1.1-2.2) and for police officers and detectives (OR: 1.8, 95%CI 1.1-2.9). Occupations where elevated risk of high grade PCa was found included gasoline station attendants (OR 4.3, 95%CI 1.8-10.4) and textile processing occupations (OR 1.8, 95%CI 1.1-3.2). Aside from logging, industries with elevated PCa risk included provincial government and financial institutions. Occupations with reduced risk included farmers (OR 0.6, 95%CI 0.4-1.0) and aircraft maintenance workers (OR 0.1, 95%CI 0.0-0.7).

Conclusions: Excess PCa risks were observed across several occupations, including predominantly white collar workers. Further analyses will focus on specific occupational exposures.

Keywords: Prostate cancer, occupation, industry

A1.2. Background

Prostate cancer (PCa) is the most frequently diagnosed cancer among Canadian men, with over 20,000 new cases and 4,000 deaths per year (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2015). PCa is also the most frequently diagnosed cancer in the United States (Siegel *et al.*, 2016) and the second worldwide after lung cancer (Ferlay *et al.*, 2015). Recognized risk factors for PCa are limited to age, first-degree family history of PCa and ethnicity (Patel and Klein, 2009). The influence of environmental and lifestyle factors in the etiology of this disease has long been suggested, including factors from the work environment, the latter having been reviewed by Parent and Siemiatycki (2001) and more recently by Doolan, et al. (2014). Notably, the International Agency for Research on Cancer (IARC) considers that limited evidence exists for arsenic and inorganic compounds, cadmium and cadmium compounds, malathion and X and gamma radiation, as well as employment in rubber manufacturing (Cogliano *et al.*, 2011b).

Many studies have focused on farming and pesticide application as elevated PCa incidence and mortality among farmers have been documented since the 1980s. Studies such as the Agricultural Health Study (AHS) in the United States have provided some clues as to agents that might be implicated, mainly organochlorines and organophosphates. More research will help characterize the potential interplay between these exposures and genetic factors in PCa etiology (Alavanja and Bonner, 2012). Other occupations that have been associated with elevated risks of PCa include firefighters (Pukkala *et al.*, 2014), aviation-related (Band *et al.*, 1999; Buja *et al.*, 2005), administrative and managerial (Sharma-Wagner *et al.*, 2000; Pukkala *et al.*, 2009), metalworking (Parent and Siemiatycki, 2001) and rubber production (Cogliano *et al.*, 2011a) occupations. However, results from occupation-based studies have generally been inconsistent or inconclusive. Very few of them have taken into account PCa aggressiveness (Potti *et al.*, 2003; Rybicki *et al.*, 2006; Su *et al.*, 2013; Su and Fontham, 2014). For instance, preliminary results from an American case-control study (Su *et al.*, 2013) found relationships between employment in truck driving or gardening occupations and aggressive PCa relative to lower PCa grade. Associations between aggressive PCa and selected organochlorine and organophosphate pesticides were also found in the AHS (Koutros *et al.*, 2013) which, in the case of Diazinon, was not apparent when looking at total PCa (Jones *et al.*, 2015).

This study aims at further exploring potential associations between employment across a wide range of occupations and industries and PCa risk, both overall and stratified by PCa grade, using data from a large, population-based case-control study conducted in Montreal, Canada.

A1.3. Methods

Study population

The Prostate Cancer and Environment Study (PROtEuS) is a population-based case-control study conducted in Greater Montreal, Canada, initiated to explore the role of lifestyle, environmental and occupational factors in PCa etiology. The study has been described in detail previously (Blanc-Lapierre *et al.*, 2015). Briefly, eligible subjects were men aged ≤75 years at diagnosis or recruitment, Canadian citizens, registered on the permanent electoral list and residing in one of the 39 electoral districts of the greater Montreal area.

Cases were all patients newly diagnosed with primary histologically confirmed PCa from September 2005 to December 2009, actively ascertained from pathology departments across the main Montreal hospitals serving the French-speaking population. These represent over 80% of all new cases in the area according to registry information. Concurrently, population controls were randomly selected from the electoral list of French-speaking men, frequency-matched to cases by 5-year age intervals. Participation rates for eligible cases and controls were 79.4% and 55.5%, respectively. Ethics boards of all participating institutions approved the study; all subjects provided written informed consent.

Data collection

Between 2006 and 2012, in-person interviews were conducted by trained interviewers, mainly at the subjects' homes, to collect a complete occupational history covering each job held for at least one year during their career, including first and last year of employment, company name and main tasks performed. Subjects also provided information on a variety of socio-demographic characteristics, anthropometric, lifestyle and environmental factors, as well as medical and residential histories. Gleason scores were extracted from cases' biopsy pathology reports to define PCa grades.

Coding of occupation and industry titles

A team of industrial hygienists reviewed the occupational histories (blind to case/control status) to assign occupation and industry titles for each job. Occupations were coded in the 1971 (updated through 1980) Canadian Classification and Dictionary of Occupations (CCDO)

(Department of Employment and Immigration, 1971) scheme, which has a hierarchical structure featuring 2-digit major groups, 3-digit minor groups, 4-digit unit groups and 7-digit occupations. Industry titles were based on the 1980 Canadian Standardized Industrial Classification (SIC) (Statistics Canada, 1980) defined by 2-digit major groups, 3-digit minor groups and 4-digit industry classes.

Statistical analyses

For each occupation and industry category, unconditional logistic regression models estimated odds ratios (OR) and 95% confidence intervals (95% CI) for the risk of PCa according to ever employment and duration of employment (<10 years and ≥10 years). For a given occupation or industry category, subjects never employed in that particular occupation or industry represented the reference group. ORs were estimated for each 2, 3, 4 and 7-digit CCDO and 2, 3 and 4-digit SIC categories with at least 10 subjects (including 1 case and 1 control) ever employed.

Models were adjusted for the three recognized risk factors for PCa: age (as a continuous variable), first-degree family history of PCa (yes, no, unknown), ancestry (European, Sub-Saharan African, Asian, Middle Eastern, other (e.g. Hispanic, Aboriginal), unknown) and timing of last PCa screening before diagnosis or interview (≤2 years, >2 years, never screened, unknown). Inclusion of the following additional covariates in the models was tested using a stepwise procedure and Akaike's Information criterion: Annual household income (<$C30,000, $C30,000-79,999, $C80,000 or more, refusal, unknown), highest level of education attained (primary, secondary/college, university, unknown), self-reported level of physical activity at work and at home (not very active, moderately active, very active), cigarette pack-years (zero, tertiles above zero, unknown), alcohol drink-years (zero, tertiles above zero, unknown) and quartiles of body-mass index (BMI) 2 years before diagnosis or interview. All variables except cigarette pack-years were retained in the final models.

In addition, multinomial logistic regression models were used to evaluate associations according to PCa grade. Low grade was defined by a Gleason score less than 7 or (3+4) and high grade by a Gleason score of 8 or higher, or (4+3) (Wright *et al.*, 2009); eight cases were excluded from this analysis due to incomplete information on Gleason scores. The same set of potential confounders retained in the binary logistic regression models was used.

Semi-Bayes adjustment for multiple comparisons

The large number of occupations and industries evaluated may lead to a non-negligible amount of false-positive associations being observed due to chance. To identify the more robust estimates, we applied a shrinkage-based method using Semi-Bayes adjustment (SB) (Greenland and Poole, 1994). Prior applications of SB methods in job title analyses have included separate adjustment for *a priori* "low" or "high" risk occupations for bladder cancer (Dryson *et al.*, 2008), or accounting for exposures to known lung carcinogens within occupations (Corbin *et al.*, 2012). In our case, there was no clear a priori expectation of occupations or industries being strongly associated with increased or decreased PCa risk. Therefore, for each combination of cancer grade (all PCas, low or high grade), type of analysis (occupations or industries), and duration (ever/never or duration in years), estimates were shrunk towards a common mean. For each analysis, a prior variance of 0.25 of the ORs on the log scale was used, corresponding to approximately 95% of the "true" ORs lying between 0.38 and 2.66 (7-fold difference). In addition, while the analysis was restricted to occupations or industries with at least 10 subjects ever employed, the stratification by duration of employment and by cancer grade could result in estimates based on a smaller sample. Therefore, the application of SB adjustment was limited to categories with at least five subjects employed.

Analyses were performed using the R software (version 3.1.0, R Development Core Team, Vienna, Austria).

A1.4. Results

The study population comprised 1937 PCa cases and 1994 population controls. Among cases, 524 were classified as having high grade PCa, 1405 low grade and 8 had insufficient information to be classified in either category. Proxy respondents, mainly spouses, completed the interview for 50 cases (3.1%) and 77 controls (3.9%). Selected characteristics of cases and controls are presented in Table I. Controls were slightly older than cases on average at interview (mean of 64.8 years) compared to cases at diagnosis (mean of 63.6 years), reflecting the slightly longer time required to secure interviews with controls. As expected, subjects from Sub-Saharan ancestry and with first-degree family history of PCa were more likely to be cases, while subjects of Asian ancestry were more likely to be controls. The proportion of controls screened at least once by prostatic specific antigen (PSA) and/or digital rectal examination (DRE) in the two years preceding the date of interview was relatively high at 76%. Annual household income, highest level of education, physical activity level, cigarette pack-years and alcohol intake level were fairly similar between cases and controls. There were more controls in the highest quartile of BMI compared to cases. The occupational histories of the 3931 subjects covered 19,373 unique jobs spanning the period 1943-2012. The number of jobs held during lifetime per subject ranged from 0 (2 cases and 2 controls) to 23, with an average of 4.9.

A total of 2993 2, 3, 4 and 7-digit CCDO groups were represented in data, with 696 (23%) having at least 10 subjects ever employed. The risk of PCa (all PCa and by PCa grade) associated with each of these 696 groups are presented in [Additional Table S1, available online at expostats.ca/jeanf/annexe1/table_s1.xls], while remaining occupations (n=2297) are listed in [Additional Table S3, available online at expostats.ca/jeanf/annexe1/table_s3.xls]. Occupations with a significantly elevated or reduced PCa risk are presented in Table II for all PCa, and in Table III for high grade PCa.

Two 2-digit categories were associated with statistically elevated ORs for ever employment in the following groups: Occupations in social sciences (OR 1.6, 95%CI: 1.1-2.2) and Forestry and logging occupations (OR 1.9, 95%CI: 1.2-3.0). Subgroups with increased PCa risk within these categories include educational counsellors (OR 3.0, 95%CI 1.1-8.3) and loggers (OR 2.0, 95%CI 1.2-3.4). Elevated ORs were also found for ever employment as Police officers and detectives (OR 1.8, 95%CI 1.1-2.9), Mixing and blending occupations (OR 3.6, 95%CI 1.2-10.8), Governmental inspectors and regulatory officers (OR 2.7, 95%CI 1.2-6.2), employment <10 years in Painting and decorating occupations excepting construction (OR 3.0, 95%CI 1.3-7.0), and in occupations associated with administration and sales, such as Receptionists, General office clerks for ever and <10 years employed, and Commodities sales clerks employed <10 years.

Associations for high grade PCa, estimated using multinomial models, were observed for ever employment in Social sciences occupations (OR 1.8, 95%CI 1.1-3.0) and Forestry and logging (OR 2.5, 95%CI 1.4-4.5). For the latter, the association was the strongest for employment ≥10 years (OR 4.4, 95%CI 1.5-12.7). Other groups where associations were mainly limited to high grade PCa were Service station attendants (OR ever 4.3, 95%CI 1.8-10.4), Textile processing occupations (OR ever 1.8, 95%CI 1.1-3.2), and Bus drivers (OR <10 years 2.9, 95%CI 1.1-7.3).

Regarding decreased PCa risk, negative associations were found for employment in Farming; specific occupations within farming where such associations were observed were General farmers (OR <10 years 0.3, 95%CI 0.1-0.9), and Field crop and Vegetable growing workers (OR ever 0.3, 95%CI 0.1-0.9). Decreased PCa risk was also found in occupations associated with Aircraft and Air transportation, and for ever and <10 years employment in Electrical engineers.

Applying SB adjustment for the analysis by occupation led to an attenuation of risk estimates, especially in occupations with few subjects ever employed. For example, the number of categories with a statistically significant association between ever employment and PCa risk saw a threefold reduction, from 37 to 12. The more robust associations for total PCa risk were found for ever employment in Occupations in social sciences (ORsb 1.5, 95%CI 1.1-2.1), Police officers (ORsb 1.7, 95%CI 1.0-2.7) and Forestry and logging occupations (ORsb 1.7, 95%CI 1.1-2.6). For high grade PCa, these included Occupations in social sciences, Forestry and logging, Bookkeeping clerks and Service station attendants. Decreased PCa risk associated with employment <10 years as Farmer also remained robust to shrinkage (ORsb 0.5, 95%CI 0.3-0.9).

For the analysis by industry title, the number of 2, 3 and 4-digit industry categories with at least one subject ever employed totaled 1125, with 613 (54%) featuring at least 10 subjects ever employed. Associations between PCa risk and employment in these 613 categories are presented in [Additional Table S2, available online at expostats.ca/jeanf/annexe1/table_s2.xls] **;** categories with at least one statistically significant association for total PCa are presented in Table IV, and in Table V for high grade PCa. The remaining 512 industry groups excluded from the analysis are listed in [Additional file 3: Table S3].

Some industries with elevated PCa risk, such as Logging industry and Protective services, provincial reflect the results from the corresponding categories for occupations, i.e. Forestry and logging occupations and Police officers and detectives, respectively. Other major industries where elevated PCa risk was found include Provincial and territorial governments (OR ever 1.5, 95%CI 1.1-2.0) and Finance, such as banks and investment intermediaries. Positive associations were also found for employment <10 years in Urban transit systems (OR 2.4, 95%CI 1.0-5.8) and ≥10 years in the Paper products industry (OR 2.3, 95%CI 1.1-5.0). Elevated risk for high grade PCa was found for Provincial and territorial governments industry (OR ever 1.7, 95%CI 1.2-2.6); in addition, elevated risk was found for employment <10 years in Local government service (OR 1.9, 95%CI 1.1-3.4). Other industries with increased risk of high grade PCa include ever employment in Wood industries (OR 1.9, 95%CI 1.0-3.6), Primary steel industries (OR 2.1, 95%CI 1.0-4.5) and Gasoline service stations (OR 2.8, 95%CI 1.4-5.5), as well as employment ≥10 years in Truck transportation (OR 2.0, 95%CI 1.1-3.5)

Lower PCa risk was found for ever employment in Agriculture (OR 0.6, 95%CI 0.4-0.9); Livestock combination farms was the only nested industry with a statistically significantly reduced association. Employments ≥10 years in the Aircraft (OR 0.5, 95%CI 0.-0.8) and Air transportation (OR 0.4, 95% 0.2-0.9) industries were associated with lower PCa risk. Employment in Food and beverage service, including restaurants, was also inversely associated with PCa.

Industries where significant positive associations with all PCa remained after SB adjustment included Logging, Chartered banks, Investment intermediaries and Provincial and territorial governments. Aside from Provincial government and Logging industries, robust associations for high grade PCa included employment ≥10 years in Truck transport (ORsb 1.7, 95%CI 1.1-2.8) and employment <10 years for Gasoline service stations (ORsb 2.1, 95%CI 1.2-3.7), Wood industries (ORsb 2.2, 95%CI 1.8, 95%CI 1.0-3.1) Local government (ORsb 1.7, 95%CI 1.0-2.8). Inverse associations for employment in Agriculture, Aircraft and Food and beverage industries remained statistically significant after SB adjustment.

A1.5. Discussion

This study explored associations between PCa risk, including high grade tumors, and employment in a wide range of occupations and industries using data from the occupational histories of approximately 4000 subjects recruited in the general population. The major substantive findings, and how they compare with results of previous studies, are detailed in the next sections, followed by an examination of the study strengths and weaknesses.

Regarding occupations and industries associated with increased PCa risk, consistently positive associations for all and high grade PCa were found for Forestry and logging occupations, with the strongest ones (OR≥1.9) found for ever employment and employment ≥10 years. Corresponding results for the Logging industry were slightly attenuated but still greater than 1.5. A recent case-control study in Northeastern Ontario found similar associations between PCa risk and employment in forestry and logging, with an OR of 2.70 (95%CI 1.21-4.79) observed for employment ≥10 years (Sritharan *et al.*, 2016). Other results in the literature range from a weakly positive association reported in a Swedish registry study (2000) to an inverse association in a large population cohort in the Nordic countries (Pukkala *et al.*, 2009). Elevated PCa risks have also been observed for other occupations related to forestry, such as forest conservationists (Alavanja *et al.*, 1989) and forest law enforcement (Sharma-Wagner *et al.*, 2000) although the associations found in our study mainly concerned logging. We also observed weaker associations for the wood industry as well as the pulp and papermaking occupations and industries. Elevated incidence or mortality from PCa associated with pulp and papermaking occupations or industries has been reported in some studies (Sharma-Wagner *et al.*, 2000; Band *et al.*, 2011). Potential exposures associated with forestry and logging include pesticides (Kangas, 2011), whole-body vibration (Jack and Oliver, 2008; Nadalin *et al.*, 2012), wood dust (Siemiatycki *et al.*, 2004), which can also be encountered in the Wood industry, and polycyclic aromatic hydrocarbons (PAHs) (Driscoll *et al.*, 2016).

Police officers and detectives represent another occupational group where increased PCa risk was found throughout our study. Most of the studies recently reviewed by Wright, et al. (2009) did not point towards elevated PCa incidence or mortality among police officers with the notable exception of a Dutch study which found a RR of 3.9 for the longest-held job (2004). Other studies not included in the Wright, et al. review found elevated PCa risk for these occupations (van der Gulden *et al.*, 1995; Finkelstein, 1998; Band *et al.*, 1999), while a previous study conducted in Montreal (Aronson *et al.*, 1996) found no association between PCa and employment for the broader Protective services occupations, which also cover firefighting. For the latter, we did observe a statistically significantly increased risk of low grade PCa for employment ≥10 years (presented in Table S1) which remained significant following SB adjustment. Potential exposures of police officers and detectives include PAH and non-ionizing radiation from radar guns (Zeegers *et al.*, 2004; Srogi, 2007; Wirth *et al.*, 2013), although the intensity of exposure associated with the use of radar guns is generally very low (IARC, 2013). These occupations may also entail night-shift work, which has been associated with prostate cancer in the literature (Rao *et al.*, 2015).

We also observed elevated PCa risk for Mixers and painting and decorating occupations, which are more directly involved with industrial chemicals. A previous case-control study in Montreal found no association between ever employment as a painter and PCa (2008) although they did not make a distinction, like us, between construction and non-construction painters. Another study within the same population did find a positive association for painting, stripping and varnishing as a leisure activity (Sharpe *et al.*, 2001). Positive associations with PCa and employment in the paint and varnish industry have also been reported in another Canadian study (Band *et al.*, 1999). These occupations/industries can be associated with a wide array of potential exposures, such as paints, lacquers, binding agents, pigments and solvents, as well as cadmium for which IARC considers to have limited evidence of an association with PCa (Cogliano *et al.*, 2011a).

There was also evidence in our study of excess PCa risk in several white collar occupations, including social sciences and administrative, management and clerical occupations. These typically entail few chemical exposures, but may reflect lower workplace physical activity levels, higher PCa screening practices, among other factors. We adjusted for these in our models with summary variables but it may be that residual confounding is at play. The literature is rather uninformative with respect to PCa risk in white collar occupations although excess risk among administrators and clerical workers has been reported (Pukkala *et al.*, 2009).

We found some occupations for which elevated risks were restricted to high grade cancer, such as Bus drivers. Excess mortality from PCa for bus drivers was reported in an American study (Krstev *et al.*, 1998). We also found elevated PCa risk for employment in the truck transport industry and, to a lesser extent, for heavy truck drivers, which have been reported in the literature, including for aggressive PCa (Järvholm and Silverman, 2003; Su and Fontham, 2014). There is some evidence that whole-body vibration might play a role in the associations between driving occupations and PCa (Young *et al.*, 2009), but other exposures such as diesel exhaust and PAHs (Aronson *et al.*, 1996; Seidler *et al.*, 1998), and circadian rhythm disruption (Mitler *et al.*, 1997) may be involved. Another group was Gasoline station attendants, where the elevated risk of high grade PCa was mainly found for employment <10 years. Previous evidence on this is limited; a study from 1987 observed a small, non-significant increase in mortality, based on three observed deaths (Schwartz, 1987). Finally, the increased risk of high-grade cancer observed for employment in occupations such as gasoline station attendants, textile processing and truck drivers does not appear to reflect delayed detection for lack of screening as associations remained essentially the same when restricting the study sample to men recently screened.

Regarding Farming and agriculture, a recent meta-analysis (Ragin *et al.*, 2013) found statistically significantly increased PCa risk among farmers, while another review (Depczynski and Lower, 2014) did not find conclusive evidence of increased PCa incidence associated with employment in farming. Our findings for these occupations were generally null or even negative for employment of less than 10 years, and are comparable to those from a previous case-control study in Montreal (Aronson *et al.*, 1996). Several factors could explain these observations. First, the study population was mainly urban, focusing on Montreal residents at recruitment, which led to a limited number of subjects previously employed in agriculture. Employment in these occupations was mainly short term (<10 years), prior to the mid-1970s, and a sizable proportion of subjects were European and Haitian immigrants. Second, employment was generally in small, family-run farms in field crop, vegetable and animal production. Chemicals thought to underlie associations with PCa in a number of previous investigations include organochlorines and organophosphates. It may well be that our study base includes fewer agricultural workers who were exposed to synthetic pesticides compared to other studies focusing on large-scale farming operations, explaining divergent findings. In addition, there is evidence that common genetic variations in xenobiotic metabolic enzymes can modulate the PCa risk associated with agricultural exposures (Koutros *et al.*, 2011), which was not accounted for in our analyses. Farmers may also be exposed to very high levels of ultraviolet radiation which may have a protective effect on PCa over several decades (Peters *et al.*, 2016). However, cumulative exposure in our study is not likely to be high due to the duration of employment of subjects ever employed as farmers was mostly less than 10 years. Finally, the generally lowered risks observed here might also reflect under-detection of PCa in our farmer group relative to other occupations, which was not fully captured through our consideration of screening practices.

Aside from farming, significantly reduced PCa risks were found for employment in aircraft and air transportation occupations or industries. Some studies have found some suggestive evidence of higher PCa risk among pilots (Krstev *et al.*, 1998; Buja *et al.*, 2005); in our case, the subjects employed within the Air transportation minor group were split between flight deck crew and support operations (e.g. air traffic controllers), resulting in small numbers within specific occupations. Aircraft maintenance workers may also perform work outdoors. However, reviews of the job descriptions indicate that most of them appear to have worked a significant portion of the time inside hangars, limiting exposure to UV radiation.

The potential associations between employment in other occupation groups and PCa risk have been explored by several studies, but findings for most specific occupations relied on a handful of studies, several of which were quite small. Moreover, large registry-based analyses focused on PCa mortality, which does not relate well to PCa incidence, and few assessed disease aggressiveness. Reviews on the subject concluded that the evidence remains scarce, and that methodological issues hampered interpretation. Some hints were apparent for occupations entailing metal and related exposures, oil-based fluids, petroleum, PAHs, engine emissions or PCBs, but there was little or no evidence accrued for occupation groups exposed to cadmium, lead, chromium and rubber products (Parent and Siemiatycki, 2001; Doolan *et al.*, 2014). The complex mixtures encountered in many occupations render them non-specific with regard to exposures per se, re-enforcing the need to conduct studies focusing on detailed exposure assessment protocols.

The job and industry titles in this study were derived from detailed information provided by subjects about their occupations, including specific tasks. Validity studies have generally shown high concordance between historical records of employment and self-reports (Teschke *et al.*, 2002). Assignment of occupational codes might have also entailed errors. However, the translation of the free-text job descriptions into standardized job and industry titles was conducted by experienced chemists-hygienists blinded to the subjects' case/control status, likely minimizing inconsistencies. The coding of job titles up to the 7-digit level of the CCDO classification also provided finer employment categories than most previous studies.

Response rates in the study were relatively high for cases (79%) and lower for controls (56%), although these values compare reasonably well with other studies based on lengthy in-person interviews. In the event that individuals who selected themselves out of the study had characteristics related to their occupation, biased risk estimates might ensue. We compared respondents and non-respondents for several socio-economic characteristics using census information, including education, income, proportion of recent immigrants and unemployment rates; for both cases and controls, differences were minimal. In addition, ORs for the associations between African or Asian ancestry and PCa, relative to European, and for first-degree family history of PCa, were comparable to those reported in the literature. These results suggest that selection bias should not be of major concern. Our study also features several important strengths, the first being the large size of the study population comprising almost 4000 subjects, making this the largest population-based case-control study investigating the role of occupational circumstances in PCa risk. Statistical power to detect associations in the more prevalent occupational groups was excellent, although it became limited for more specific employment categories.

Second, it relied on histologically confirmed PCa cases; for controls, screening rates were very high with over 75% reporting having been screened by PSA and/or DRE within the 2 years preceding the date of interview. In Montreal, Canada, access to the medical system is universal and free of charge. While there is no systematic PCa screening program in place, at the time of study screening was frequently integrated within the annual routine medical examination, independently of socio-economic status. Screening behavior can be influenced by a number of factors, including lifestyle, beliefs, and medical follow-ups, including those offered in the workplace, and thus can potentially confound an occupation-PCa association. The available information on PCa screening patterns in our population allowed us to adjust for screening in the models and limit the impact of disease misclassification on our results and conclusions. Moreover, analyses excluding controls not screened for PCa in the previous two years to limit the potential for latent PCa yielded similar findings.

Third, information on Gleason scores for cases was extracted from pathology reports and was available for nearly all subjects, allowing us to investigate potential associations stratified by PCa grade; this has generally been overlooked in studies of occupational risk factors for PCa. Fourth, the information available on recognized risk factors (e.g. ancestry, family history) and other factors (e.g. BMI, education level) associated directly or indirectly with PCa allowed us to account for these potential confounders in the inferences made.

Lastly, we used statistical methods to account for the large number of associations between employment by occupation or industry and PCa risk in order to identify the most robust results warranting further investigation. On average, 30% of the statistically significant ORs remained as such following SB adjustment across the different analyses performed.

Little data has been accrued to date on the role of specific occupational exposures in prostate cancer risk. Agents for which there is limited evidence for humans include arsenic and inorganic arsenic compounds, cadmium and cadmium compounds, malathion, among others (Cogliano *et al.*, 2011a), and a role for night work is also suspected (Rao *et al.*, 2015). Some of these may or may not explain some of the associations observed here, but in the present population-based analysis, job and industry titles should not be viewed as reliable proxies for specific occupational exposures. Nevertheless, emerging associations can serve as leads for future work investigating specific exposures. On one hand, the identification of positive associations between employment in a given occupation or industry and PCa risk could be due to exposure to one agent, or to a diversity of agents and mixtures, not discounting the possibility of interactions between agents. On the other hand, exposure to an agent associated with PCa might occur in a range of occupations and the association might not be picked up with an analysis based on job titles, especially if the prevalence of exposure within occupations is low. Grouping subjects according to specific exposures, rather than job titles, is thus a more powerful approach to detect associations (Siemiatycki *et al.*, 1981). Indeed, one of the limitations identified in reviews of occupational risk factors for PCa (Parent and Siemiatycki, 2001; Doolan *et al.*, 2014) has centered on the use of crude methods for exposure assessment in many studies. To this end, the translation of the occupational histories in the PROtEuS study into estimates of exposure to over 300 agents is currently ongoing.

A1.6. Conclusions

Our findings suggest excess PCa risks among forestry workers and policemen, as well as in some predominantly white collar occupations and industries such as public service, administrative and clerical work. This represents an interesting research lead as little evidence has been accrued to date on white collar occupations and cancer risk. Increased risk of high grade PCa was also found for other occupations such as gas station attendants and bus drivers that were not apparent when looking at low and high PCa grades combined. Additional investigations are needed to identify specific exposures or circumstances potentially associated with PCa among subjects employed in these occupations.

Additional files

Additional file 1: Table S1. Associations between PCa risk (all PCa, low-grade and high-grade PCa) for ever employment and duration of employment, occupations (XLS 703 kb)

Additional file 2: Table S2. Associations between PCa risk (all PCa, low-grade and high-grade PCa) for ever employment and duration of employment, industries (XLS 601 kb)

Additional file 3: Table S3. List of occupation and industry categories featuring at least one subject ever employed and <10 subjects ever employed or 0 case or control (XLS 450 kb)

A1.7. Declarations

List of abbreviations

AHS: (Agricultural Health Study); BMI (Body-mass index); CCDO: (Canadian Classification and Dictionary of Occupations); CI: (Confidence interval); n.e.c.: (not elsewhere classified); OR: (Odds ratio); PAH: (Polycyclic aromatic hydrocarbon); PCa: (Prostate cancer); PSA: (Prostatic specific antigen); SB: (Semi-Bayes); SIC: (Standardized Industrial Classification)

Acknowledgments

The contribution of the entire Epidemiology and Biostatistics unit at INRS-Institut Armand-Frappier is gratefully acknowledged. In particular, we want to highlight the work of Robert Bourbonnais, Ramzan Lakhani, Louise Nadon, Dora Rodriguez and Mounia Rhazi in reviewing questionnaire data and coding of occupation and industry titles. We thank the urologists from the participating hospitals for their collaboration.

Funding

Funding for this study was supported by the Canadian Cancer Society (grants no. 13149, 19500, 19864, 19865), the Cancer Research Society, the Fonds de Recherche du Québec – Santé (FRQS), FRQS-RRSE and the Ministère du Dévelppement Économique, de l’Innovation et de l’Exportation du Québec. JFS was supported by the FRQS and the Institut de Recherche Robert-Sauvé en Santé et en Sécurité du Travail. MEP and JL hold career awards from the FRQS. Analyses were conducted on the supercomputer Briarée which is funded by the Canada Foundation for Innovation, NanoQuébec, RMGA and the Fonds de recherche du Québec—Nature et technologies (FRQ-NT).

Availability of data and materials

The data used for this project is not shared on public repositories for confidentiality reasons and since the analysis of the PROtEuS study is ongoing. The data can be made available for use in research collaborations by contacting the corresponding author of this manuscript.

Authors' contributions

JFS conducted the data analysis and preparation of the paper. MEP contributed to the conception, design and data acquisition of the PROtEuS study. JFS, MEP and JL contributed to the interpretation of the data and critical revision of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests

Consent for publication

Not applicable.

Ethics approval and consent to participate

This study was approved by the Ethics Committees of the following institutions: Institut national de la recherche scientifique, Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Hôpital Maisonneuve-Rosemont, Hôpital Jean- Talon, Hôpital Fleury, and Hôpital Charles-LeMoyne. All participants provided written informed consent.

A1.8. Tables and figures

Table I. Selected characteristics of cases and controls

Variable	Controls (n=1994)	Cases: All (n=1937)	Cases: Low grade [1] (n=1405)	Cases: High grade [2] (n=524)
Age at diagnosis or interview in years, n (%)				
<60	446 (22.4)	523 (27.0)	414 (29.5)	107 (20.4)
≥60 and <65	450 (22.6)	486 (25.1)	351 (25.0)	131 (25.0)
≥65 and <70	522 (26.2)	498 (25.7)	347 (24.7)	150 (28.6)
≥70 and ≤75	576 (28.9)	430 (22.2)	293 (20.9)	136 (26.0)
Ancestry, n (%)				
European	1685 (84.5)	1696 (87.6)	1234 (87.8)	455 (86.8)
Sub-Saharan	90 (4.5)	130 (6.7)	99 (7.0)	30 (5.7)
Asian	73 (3.7)	24 (1.2)	15 (1.1)	9 (1.7)
Greater Middle East	99 (5.0)	45 (2.3)	32 (2.3)	13 (2.5)
Other	33 (1.7)	30 (1.5)	18 (1.3)	12 (2.3)
Don't Know	14 (0.7)	12 (0.6)	7 (0.5)	5 (1.0)
First-degree family history of PCa, n (%)				
No	1739 (87.2)	1419 (73.3)	1008 (71.7)	405 (77.3)
Yes	199 (10.0)	452 (23.3)	351 (25.0)	100 (19.1)
Don't Know	56 (2.8)	66 (3.4)	46 (3.3)	19 (3.6)
Date of last screening for PCa before diagnosis or interview, n (%)				
≤2 years	1510 (75.7)	1917 (99.0)	1388 (98.8)	521 (99.4)
> 2 years	235 (11.8)	1 (0.1)	1 (0.1)	0 (0.0)
No/Never	191 (9.6)	3 (0.2)	2 (0.1)	1 (0.2)
Don't Know	58 (2.9)	16 (0.8)	14 (1.0)	2 (0.4)
Annual household income in CAD$, n (%)				
<10,000-29,999$	497 (24.9)	490 (25.3)	323 (23.0)	167 (31.9)
30,000-79,999$	872 (43.8)	874 (45.1)	640 (45.6)	228 (43.5)
80,000 and more	428 (21.5)	426 (22.0)	343 (24.4)	81 (15.5)
Prefers not to respond	186 (9.3)	132 (6.8)	91 (6.5)	41 (7.8)
Don't Know	9 (0.5)	15 (0.8)	8 (0.6)	7 (1.3)
Highest level of education attained, n (%)				
Primary	428 (21.5)	449 (23.2)	312 (22.2)	136 (26.0)
Secondary/College	953 (47.8)	891 (46.0)	629 (44.8)	259 (49.5)
University	611 (30.7)	592 (30.6)	461 (32.8)	127 (24.3)
Don't Know	1 (0.1)	4 (0.2)	3 (0.2)	1 (0.2)

Variable	Controls (n=1994)	Cases All (n=1937)	Cases Low grade [1] (n=1405)	Cases High grade [2] (n=524)
Self-reported level of physical activity at work or leisure, n (%)				
Not very active	151 (7.6)	127 (6.6)	84 (6.0)	43 (8.2)
Moderately active	753 (37.8)	682 (35.2)	516 (36.7)	160 (30.5)
Very active	1089 (54.6)	1128 (58.2)	805 (57.3)	321 (61.3)
Cigarette pack-years, n (%)				
Zero	544 (27.3)	542 (28.0)	404 (28.8)	136 (26.0)
>0 - 15.1	441 (22.1)	498 (25.7)	381 (27.1)	114 (21.8)
>15.1 - 39.4	524 (26.3)	420 (21.7)	296 (21.1)	122 (23.3)
>39.4 - 223	474 (23.8)	458 (23.6)	313 (22.3)	144 (27.5)
Don't know	11 (0.6)	19 (1.0)	11 (0.8)	8 (1.5)
Alcohol drink-years, n (%)				
Zero	231 (11.6)	214 (11.0)	157 (11.2)	56 (10.7)
> 0 - 24.4	575 (28.8)	548 (28.3)	413 (29.4)	133 (25.4)
> 24.4 - 76	569 (28.5)	553 (28.5)	408 (29.0)	143 (27.3)
> 76 - 2660	577 (28.9)	545 (28.1)	376 (26.8)	166 (31.7)
Don't know	42 (2.1)	77 (4.0)	51 (3.6)	26 (5.0)
Body-mass index, 2 years before diagnosis or interview, in kg/m^2, n (%)				
≤ 24.2	477 (23.9)	502 (25.9)	362 (25.8)	137 (26.2)
> 24.2 - 26.5	492 (24.7)	512 (26.4)	373 (26.5)	137 (26.2)
> 26.5 - 29.2	477 (23.9)	479 (24.7)	362 (25.8)	116 (22.2)
> 29.2	535 (26.8)	431 (22.3)	299 (21.3)	130 (24.9)
Don't know	13 (0.7)	12 (0.6)	9 (0.6)	3 (0.6)

1. Gleason score ≤6 or 3+4
2. Gleason score ≥8 or 4+3

Table II. Selected associations between selected occupations and risk of prostate cancer (all cancers)

	Never	Ever employed			<10 years employed			≥10 years employed		
CCDO code and description	**Ca/Co**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**
111: Officials and administrators unique to government	1875/1936	62/58	1.2 (0.8-1.8)	1.2 (0.8-1.7)	19/31	0.7 (0.4-1.3)	0.8 (0.5-1.3)	43/27	1.8 (1.0-3.1)*	1.6 (0.9-2.5)
1116: Inspectors and regulatory officers, government	1910/1982	27/12	2.7 (1.2-6.2)*	1.8 (1.0-3.4)	11/4	2.9 (0.8-10.7)	1.5 (0.7-3.3)	16/8	2.6 (0.9-7.4)	1.6 (0.8-3.2)
1130126: General manager, finance (bank. & finance)	1925/1989	12/5	3.5 (1.0-12.5)*	1.6 (0.8-3.5)	5/3	1.9 (0.4-10.2)	1.2 (0.5-2.8)	7/2	6.7 (0.9-48.2)	1.5 (0.6-3.5)
1143: Production management occupations	1911/1954	26/40	0.7 (0.4-1.3)	0.8 (0.5-1.3)	12/12	1.6 (0.6-4.0)	1.3 (0.7-2.5)	14/28	0.5 (0.2-1.0)*	0.6 (0.4-1.1)
1171162: Auditor (prof. & tech., n.e.c.)	1908/1975	29/19	1.7 (0.9-3.3)	1.4 (0.8-2.5)	9/11	0.9 (0.3-2.3)	0.9 (0.5-1.9)	20/8	2.9 (1.1-7.3)*	1.8 (0.9-3.4)
1179299: Other occupations related to management and administration	1916/1985	21/9	2.8 (1.1-6.9)*	1.7 (0.9-3.4)	12/6	2.5 (0.8-7.7)	1.5 (0.7-3.1)	9/3	3.4 (0.7-16.3)	1.4 (0.6-3.3)
21: Occupations in natural sciences, engineering and mathematics	1713/1722	224/272	0.8 (0.7-1.1)	0.8 (0.7-1.1)	85/75	1.1 (0.8-1.6)	1.1 (0.8-1.6)	139/197	0.7 (0.6-0.9)*	0.8 (0.6-1.0)*
2144: Electrical engineers	1926/1965	11/29	0.4 (0.2-0.8)*	0.5 (0.3-1.0)*	2/13	0.1 (0.0-0.5)*	0.5 (0.2-1.2)	9/16	0.8 (0.3-1.9)	0.9 (0.4-1.7)
2144110: Design and development engineer, electrical and electronic (prof. & tech., n.e.c.)	1934/1987	3/7	0.2 (0.1-1.0)	0.6 (0.3-1.5)	1/5	0.1 (0.0-0.8)*	0.7 (0.3-1.6)	2/2	1.5 (0.1-17.6)	
23: Occupations in social sciences and related fields	1809/1914	128/80	1.6 (1.1-2.2)*	1.5 (1.1-2.1)*	52/30	1.6 (1.0-2.6)	1.4 (0.9-2.3)	76/50	1.6 (1.0-2.5)*	1.5 (1.0-2.2)*
2391118: Counsellor, educational (educ.)	1918/1988	19/6	3.0 (1.1-8.3)*	1.7 (0.9-3.5)	12/4	2.4 (0.8-7.4)	1.4 (0.7-3.0)	7/2	5.7 (0.8-43.4)	1.4 (0.6-3.4)
271: University teaching and related occupations	1852/1933	85/61	1.5 (1.0-2.2)*	1.4 (1.0-2.0)	29/20	1.4 (0.7-2.6)	1.3 (0.7-2.2)	56/41	1.5 (1.0-2.4)	1.4 (0.9-2.2)
315: Other occupations in medicine and health	1913/1973	24/21	1.0 (0.5-1.9)	1.0 (0.6-1.7)	19/6	3.2 (1.0-9.8)*	1.7 (0.8-3.5)	5/15	0.3 (0.1-0.9)*	0.6 (0.3-1.2)
335: Occupations in writing	1919/1962	18/32	0.6 (0.3-1.2)	0.7 (0.4-1.2)	10/12	1.5 (0.5-4.3)	1.2 (0.6-2.5)	8/20	0.3 (0.1-0.8)*	0.6 (0.3-1.0)
4171: Receptionists and information clerks	1918/1986	19/8	3.0 (1.1-8.0)*	1.7 (0.9-3.5)	19/8	3.0 (1.1-8.0)*	1.7 (0.9-3.5)	0/0		
4197: General office clerks	1839/1910	98/84	1.3 (0.9-1.8)	1.2 (0.9-1.7)	81/53	1.6 (1.1-2.3)*	1.5 (1.0-2.1)*	17/31	0.6 (0.3-1.2)	0.7 (0.4-1.3)
513/514: Sales occupations, commodities	1564/1569	373/425	0.9 (0.7-1.0)	0.9 (0.7-1.0)	198/193	1.0 (0.8-1.3)	1.0 (0.8-1.3)	175/232	0.8 (0.6-0.9)*	0.8 (0.6-0.9)*
5135182: Salesperson, footwear (ret. Trade)	1936/1985	1/9	0.1 (0.0-0.9)*	0.7 (0.3-1.7)	0/7			1/2	0.8 (0.1-8.9)	
5137: Sales clerks, commodities	1833/1908	104/86	1.2 (0.9-1.7)	1.2 (0.9-1.6)	82/58	1.5 (1.0-2.3)*	1.5 (1.0-2.1)*	22/28	0.7 (0.4-1.2)	0.8 (0.4-1.2)
5137111: Supermarket clerk (ret. Trade)	1900/1968	37/26	1.6 (0.9-2.9)	1.4 (0.9-2.4)	32/18	2.3 (1.2-4.7)*	1.8 (1.0-3.1)*	5/8	0.5 (0.2-1.7)	0.8 (0.4-1.6)
5145: Service station attendants	1912/1982	25/12	2.3 (1.0-5.1)*	1.7 (0.9-3.1)	21/11	2.0 (0.9-4.8)	1.5 (0.8-2.9)	4/1	4.6 (0.5-42.0)	1.3 (0.5-3.2)
5199158: Telephone solicitor (any ind.)	1930/1981	7/13	0.5 (0.2-1.3)	0.7 (0.3-1.4)	3/12	0.2 (0.1-0.9)*	0.6 (0.3-1.3)	4/1	2.7 (0.3-24.6)	1.2 (0.5-2.9)
6112: Police officers and detectives, government	1892/1964	45/30	1.8 (1.1-2.9)*	1.6 (1.0-2.4)	13/5	3.2 (1.0-10.0)*	1.7 (0.8-3.5)	32/25	1.5 (0.8-2.6)	1.4 (0.8-2.2)

	Never	Ever employed			<10 years employed			≥10 years employed		
CCDO code and description	**Ca/Co**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**
6112146: Detective (gov. Serv.)	1925/1987	12/7	2.0 (0.7-5.2)	1.4 (0.7-2.8)	9/1	10.3 (1.3-82.9)*	1.5 (0.6-3.8)	3/6	0.5 (0.1-2.3)	0.8 (0.4-1.9)
6112158: Police officer (gov. Serv.)	1899/1972	38/22	2.0 (1.1-3.5)*	1.7 (1.0-2.7)*	18/7	3.4 (1.3-8.8)*	1.9 (0.9-3.7)	20/15	1.4 (0.7-2.9)	1.3 (0.7-2.3)
6117190: Infantry soldier (military)	1877/1898	60/96	0.7 (0.5-1.0)	0.7 (0.5-1.0)	57/95	0.7 (0.5-1.0)*	0.7 (0.5-1.0)	3/1	2.4 (0.2-24.2)	
612: Food and beverage preparation and related service occupations	1797/1813	140/181	0.7 (0.6-1.0)*	0.8 (0.6-1.0)*	77/103	0.7 (0.5-1.0)*	0.7 (0.5-1.0)	63/78	0.8 (0.6-1.2)	0.8 (0.6-1.2)
6121130: Short-order cook (cater. & lodg.)	1933/1982	4/12	0.2 (0.1-0.8)*	0.6 (0.3-1.3)	2/7	0.3 (0.0-1.6)	0.8 (0.3-1.8)	2/5	0.2 (0.0-1.2)	0.7 (0.3-1.7)
613: Occupations in lodging and other accommodation	1915/1965	22/29	0.7 (0.4-1.2)	0.8 (0.4-1.3)	16/13	1.1 (0.5-2.5)	1.1 (0.6-2.0)	6/16	0.3 (0.1-0.9)*	0.6 (0.3-1.2)
6145: Travel and related attendants, except food and beverage	1935/1986	2/8	0.2 (0.0-0.9)*	0.6 (0.3-1.5)	1/4	0.2 (0.0-2.4)	0.8 (0.3-2.0)	1/4	0.1 (0.0-1.5)	0.8 (0.3-1.9)
619: Other service occupations	1759/1799	178/195	0.8 (0.6-1.0)*	0.8 (0.6-1.0)*	117/113	0.8 (0.6-1.1)	0.8 (0.6-1.1)	61/82	0.7 (0.5-1.0)	0.7 (0.5-1.0)
6191126: Hospital cleaner (misc. Serv.)	1921/1973	16/21	0.6 (0.3-1.2)	0.7 (0.4-1.3)	8/15	0.3 (0.1-0.9)*	0.6 (0.3-1.1)	8/6	1.4 (0.4-4.6)	1.2 (0.6-2.5)
6198: Occupations in labouring and other elemental work, services	1871/1931	66/63	0.9 (0.6-1.3)	0.9 (0.6-1.3)	60/47	1.0 (0.7-1.6)	1.0 (0.7-1.6)	6/16	0.3 (0.1-0.9)*	0.6 (0.3-1.2)
71: Farming. Horticultural and animal-husbandry occupations	1840/1860	97/134	0.7 (0.5-1.0)*	0.8 (0.6-1.0)	59/82	0.7 (0.5-1.1)	0.8 (0.5-1.1)	38/52	0.7 (0.4-1.1)	0.8 (0.5-1.2)
711: Farmers	1903/1941	34/53	0.6 (0.4-1.0)*	0.7 (0.4-1.0)	14/31	0.4 (0.2-0.8)*	0.5 (0.3-0.9)*	20/22	0.9 (0.5-1.9)	1.0 (0.5-1.7)
7111: General farmers	1923/1974	14/20	0.6 (0.3-1.2)	0.7 (0.4-1.3)	5/12	0.3 (0.1-0.9)*	0.6 (0.3-1.2)	9/8	1.1 (0.4-3.2)	1.0 (0.5-2.1)
7111110: Farmer, general (agric.)	1923/1974	14/20	0.6 (0.3-1.2)	0.7 (0.4-1.3)	5/12	0.3 (0.1-0.9)*	0.6 (0.3-1.2)	9/8	1.1 (0.4-3.2)	1.0 (0.5-2.1)
7183: Field crop and vegetable-growing workers	1932/1978	5/16	0.3 (0.1-0.9)*	0.6 (0.3-1.2)	4/11	0.4 (0.1-1.4)	0.7 (0.3-1.6)	1/5	0.2 (0.0-1.6)	0.8 (0.3-1.9)
7183122: Farm worker, vegetable (agric.)	1935/1983	2/11	0.2 (0.0-0.9)*	0.6 (0.3-1.5)	2/7	0.4 (0.1-2.4)	0.8 (0.3-1.9)	0/4		
75: Forestry and logging occupations	1874/1964	63/30	1.9 (1.2-3.0)*	1.7 (1.1-2.6)*	46/23	1.7 (1.0-3.0)	1.5 (0.9-2.5)	17/7	2.4 (0.9-6.2)	1.6 (0.8-3.1)
751: Forestry and logging occupations	1874/1964	63/30	1.9 (1.2-3.0)*	1.7 (1.1-2.6)*	46/23	1.7 (1.0-3.0)	1.5 (0.9-2.5)	17/7	2.4 (0.9-6.2)	1.6 (0.8-3.1)
7513: Timber cutting and related occupations	1884/1969	53/25	1.9 (1.1-3.2)*	1.6 (1.0-2.6)*	38/20	1.7 (0.9-3.0)	1.5 (0.9-2.4)	15/5	2.8 (0.9-8.4)	1.6 (0.8-3.3)
7513122: Logger, all-round (forest. & log.)	1884/1970	53/24	2.0 (1.2-3.4)*	1.7 (1.1-2.7)*	38/19	1.8 (1.0-3.2)	1.5 (0.9-2.5)	15/5	2.8 (0.9-8.4)	1.6 (0.8-3.3)
8161: Mixing and blending occupations, chemicals and related materials	1919/1986	18/8	3.6 (1.2-10.8)*	1.8 (0.9-3.7)	12/4	8.0 (1.3-50.2)*	1.6 (0.7-3.8)	6/4	1.9 (0.4-7.7)	1.2 (0.6-2.8)
8161218: Mixer (chem., n.e.c.; paint & varn.)	1926/1991	11/3	8.0 (1.2-53.8)*	1.6 (0.6-3.7)	8/2	6.2 (0.8-47.8)	1.4 (0.6-3.5)	3/1	20.6 (0.22.7)	
8529: Other fabricating and assembling occupations, metal products, n.e.c.	1922/1986	15/8	1.8 (0.6-4.9)	1.3 (0.6-2.7)	8/7	0.8 (0.2-2.5)	0.9 (0.4-1.9)	7/1	19.8 (1.0.5)*	1.4 (0.5-3.4)
858: Mechanics and repairmen, n.e.c.	1782/1811	155/183	0.8 (0.6-1.1)	0.8 (0.6-1.1)	63/64	1.1 (0.7-1.6)	1.1 (0.7-1.6)	92/119	0.7 (0.5-1.0)*	0.7 (0.5-1.0)*
8581118: Industrial-truck mechanic (mech. Equip., n.e.c.)	1934/1983	3/11	0.2 (0.0-0.8)*	0.6 (0.3-1.3)	1/4	0.4 (0.0-4.2)	0.9 (0.4-2.2)	2/7	0.1 (0.0-0.8)*	0.6 (0.3-1.4)

	Never	Ever employed			<10 years employed			≥10 years employed		
CCDO code and description	**Ca/Co**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**
8582: Aircraft mechanics and repairmen	1936/1976	1/18	0.1 (0.0-0.7)*	0.6 (0.3-1.6)	0/6			1/12	0.1 (0.0-1.1)	0.7 (0.3-1.7)
8595: Painting and decorating occupations, except construction	1902/1975	35/19	1.9 (1.0-3.5)	1.6 (0.9-2.6)	27/9	3.0 (1.3-7.0)*	1.9 (1.0-3.6)	8/10	0.8 (0.3-2.4)	0.9 (0.4-1.9)
873: Electrical power, lighting and wire communications equipment erecting, installing and repairing occupations	1852/1924	85/70	1.3 (0.9-1.8)	1.2 (0.9-1.8)	41/28	2.0 (1.1-3.6)*	1.7 (1.0-2.8)*	44/42	0.9 (0.6-1.5)	1.0 (0.6-1.5)
8739: Electrical power, lighting and wire communications equipment erecting, installing and repairing occupations, n.e.c.	1933/1980	4/14	0.3 (0.1-0.8)*	0.6 (0.3-1.2)	2/8	0.2 (0.0-1.1)	0.7 (0.3-1.6)	2/6	0.3 (0.1-1.6)	0.7 (0.3-1.7)
8784: Plasterers and related occupations	1933/1985	4/9	0.2 (0.1-0.9)*	0.6 (0.3-1.3)	0/5			4/4	0.4 (0.1-1.8)	0.8 (0.3-1.7)
8791: Pipefitting, plumbing and related occupations, n.e.c.	1886/1954	51/40	1.7 (1.0-2.8)*	1.5 (1.0-2.4)	28/19	1.8 (0.9-3.5)	1.5 (0.8-2.6)	23/21	1.6 (0.7-3.4)	1.3 (0.7-2.4)
8799: Other construction trades occupations, n.e.c.	1877/1948	60/46	1.8 (1.1-2.9)*	1.6 (1.0-2.5)*	36/26	1.6 (0.9-3.0)	1.4 (0.8-2.4)	24/20	2.1 (0.9-4.6)	1.6 (0.8-2.9)
911: Air transport operating operations	1922/1974	15/20	0.6 (0.3-1.3)	0.7 (0.4-1.3)	11/9	1.3 (0.5-3.6)	1.2 (0.6-2.3)	4/11	0.2 (0.1-0.7)*	0.6 (0.3-1.2)
915: Water transport operating occupations	1923/1971	14/23	0.6 (0.3-1.2)	0.7 (0.4-1.3)	7/18	0.3 (0.1-0.8)*	0.6 (0.3-1.1)	7/5	1.7 (0.5-6.3)	1.2 (0.6-2.7)
9173: Taxi drivers and chauffeurs	1870/1937	67/57	1.2 (0.8-1.8)	1.2 (0.8-1.7)	24/29	0.8 (0.4-1.4)	0.8 (0.5-1.4)	43/28	1.8 (1.0-3.2)*	1.6 (0.9-2.6)
9175129: Solid waste collection truck driver (motor trans.)	1935/1985	2/9	0.2 (0.0-0.9)*	0.6 (0.3-1.5)	2/7	0.2 (0.0-1.2)	0.7 (0.3-1.6)	0/2		
9179118: Dispatcher, motor vehicles (motor trans.)	1925/1991	12/3	8.4 (1.2-57.6)*	1.6 (0.6-3.8)	9/2	6.8 (0.9-53.9)	1.4 (0.6-3.5)	3/1	18.6 (0.22.5)	
9918: Occupations in labouring and other elemental work, n.e.c.	1920/1982	17/12	2.7 (0.9-7.9)	1.6 (0.8-3.3)	12/7	4.1 (1.0-16.6)*	1.6 (0.7-3.6)	5/5	1.3 (0.2-6.7)	1.1 (0.5-2.5)
9918110: Labourer, municipal (gov. Serv.)	1920/1983	17/11	3.4 (1.0-10.8)*	1.7 (0.8-3.5)	12/6	6.5 (1.2-33.5)*	1.6 (0.7-3.8)	5/5	1.3 (0.2-6.7)	1.1 (0.5-2.5)

§ Occupations selected had at least one significantly elevated or reduced association with prostate cancer among duration of employment categories (ever, <10 years or ≥10 years). Odds ratios adjusted for age, first-degree family history of prostate cancer, ancestry, screening for prostate cancer, annual household income, highest level of education attained, level of physical activity, alcohol intake and body mass index. * The 95% confidence interval excludes the null value. n.e.c.: not elsewhere classified.

Table III. Selected associations between selected occupations and risk of high grade prostate cancer

	Never	Ever employed			<10 years employed			≥10 years employed		
CCDO code and description	**Ca/Co**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**
1135122: Credit manager (prof. & tech., n.e.c.)	520/1990	4/4	4.0 (1.0-17.0)	1.6 (0.7-3.7)	2/1	12.6 (1.1.4)*		2/3	2.2 (0.3-13.8)	1.3 (0.6-3.1)
1137: Sales and advertising management occupations	505/1931	19/63	1.3 (0.7-2.2)	1.2 (0.8-2.0)	12/22	2.1 (1.0-4.5)*	1.7 (0.9-3.1)	7/41	0.7 (0.3-1.7)	0.9 (0.5-1.7)
1137118: Manager, sales (prof. & tech., n.e.c.)	508/1945	16/49	1.4 (0.8-2.6)	1.3 (0.8-2.2)	10/17	2.8 (1.2-6.5)*	1.9 (1.0-3.6)*	6/32	0.7 (0.3-1.8)	0.9 (0.5-1.8)
2163: Draughtsmen	512/1949	12/45	1.2 (0.6-2.4)	1.2 (0.7-2.0)	8/15	2.6 (1.0-6.6)*	1.8 (0.9-3.4)	4/30	0.6 (0.2-1.7)	0.8 (0.4-1.7)
23: Occupations in social sciences and related fields	494/1914	30/80	1.8 (1.1-3.0)*	1.7 (1.1-2.6)*	17/30	2.4 (1.3-4.6)*	1.9 (1.1-3.3)*	13/50	1.4 (0.7-2.7)	1.3 (0.8-2.3)
2349: Occupations in law and jurisprudence, n.e.c.	520/1990	4/4	4.3 (0.9-19.9)	1.6 (0.7-3.7)	4/1	12.2 (1.3.2)*	1.7 (0.7-4.1)	0/3		
2711199: Other university teachers	520/1992	4/2	7.0 (1.2-39.6)*	1.7 (0.7-4.0)	3/2	4.8 (0.8-29.4)	1.6 (0.7-3.7)	1/0		
279: Other teaching and related occupations	501/1933	23/61	1.7 (1.0-2.8)	1.5 (1.0-2.4)	10/38	1.2 (0.6-2.6)	1.2 (0.7-2.2)	13/23	2.3 (1.1-4.8)*	1.8 (1.0-3.2)
4131134: Accounting clerk (clerical)	514/1976	10/18	2.8 (1.2-6.8)*	1.8 (1.0-3.5)	7/14	2.4 (0.9-6.5)	1.6 (0.8-3.3)	3/4	4.8 (0.8-28.6)	1.6 (0.7-3.7)
4131142: Bookkeeping clerk (clerical)	513/1973	11/21	2.8 (1.2-6.3)*	1.9 (1.0-3.6)*	9/15	3.3 (1.3-8.5)*	2.0 (1.0-3.9)*	2/6	1.6 (0.3-9.1)	1.2 (0.5-2.9)
4133110: Teller (bank. & finance)	512/1969	12/25	2.1 (1.0-4.4)	1.6 (0.9-3.0)	12/24	2.1 (1.0-4.4)*	1.7 (0.9-3.0)	0/1		
5133114: Pharmaceutical representative (whole. Trade)	520/1986	4/8	1.9 (0.6-6.5)	1.3 (0.6-2.9)	4/4	4.6 (1.1-19.3)*	1.8 (0.8-4.0)	0/4		
5135178: Salesperson, wearing apparel (ret. Trade; whole. Trade)	514/1980	10/14	3.2 (1.3-7.6)*	2.0 (1.0-3.8)*	7/10	3.0 (1.1-8.3)*	1.8 (0.9-3.6)	3/4	3.8 (0.7-19.4)	1.6 (0.7-3.6)
5145: Service station attendants	511/1982	13/12	4.3 (1.8-10.4)*	2.4 (1.2-4.5)*	11/11	4.0 (1.6-10.2)*	2.2 (1.1-4.3)*	2/1	7.0 (0.6-78.7)	
5145110: Service-station attendant (motor vehicle; ret. Trade)	511/1982	13/12	4.4 (1.8-10.5)*	2.4 (1.2-4.5)*	11/11	4.0 (1.6-10.3)*	2.2 (1.1-4.3)*	2/1	7.5 (0.7-84.2)	
5172: Real estate salesmen	508/1960	16/34	1.8 (0.9-3.3)	1.5 (0.9-2.6)	10/16	2.4 (1.0-5.5)*	1.8 (0.9-3.3)	6/18	1.2 (0.5-3.3)	1.2 (0.6-2.4)
5172118: Salesperson, real estate (insur. & real estate)	508/1965	16/29	2.0 (1.1-3.9)*	1.7 (1.0-2.9)	10/14	2.9 (1.2-6.8)*	1.9 (1.0-3.7)	6/15	1.4 (0.5-3.7)	1.2 (0.6-2.5)
6112: Police officers and detectives, government	513/1964	11/30	1.4 (0.7-2.9)	1.3 (0.7-2.3)	5/5	4.0 (1.0-15.1)*	1.8 (0.8-3.9)	6/25	0.9 (0.4-2.3)	1.0 (0.5-2.0)
75: Forestry and logging occupations	500/1964	24/30	2.5 (1.4-4.5)*	2.0 (1.2-3.3)*	15/23	2.0 (1.0-4.0)	1.6 (0.9-2.9)	9/7	4.4 (1.5-12.7)*	2.1 (1.0-4.3)*
751: Forestry and logging occupations	500/1964	24/30	2.5 (1.4-4.5)*	2.0 (1.2-3.3)*	15/23	2.0 (1.0-4.0)	1.6 (0.9-2.9)	9/7	4.4 (1.5-12.7)*	2.1 (1.0-4.3)*
7513: Timber cutting and related occupations	504/1969	20/25	2.5 (1.3-4.7)*	1.9 (1.1-3.3)*	12/20	1.8 (0.9-4.0)	1.5 (0.8-2.8)	8/5	5.2 (1.5-17.6)*	2.1 (1.0-4.4)
7513122: Logger, all-round (forest. & log.)	504/1970	20/24	2.6 (1.4-5.0)*	2.0 (1.2-3.4)*	12/19	2.0 (0.9-4.2)	1.6 (0.9-2.9)	8/5	5.2 (1.5-17.6)*	2.1 (1.0-4.4)

	Never	Ever employed			<10 years employed			≥10 years employed		
CCDO code and description	**Ca/Co**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**
8131: Metal smelting, converting and refining furnacemen	522/1991	2/3	2.3 (0.3-16.8)	1.3 (0.5-3.0)	1/2	1.1 (0.1-12.7)		1/1	53.7 (13.2.3)*	
821/822: Food, beverage and related processing occupations	503/1882	21/112	0.5 (0.3-0.9)*	0.6 (0.4-1.0)*	12/74	0.5 (0.2-0.9)*	0.6 (0.4-1.1)	9/38	0.6 (0.3-1.4)	0.8 (0.4-1.5)
825: Pulp and papermaking and related occupations	520/1991	4/3	6.3 (1.1-35.6)*	1.7 (0.7-3.9)	3/2	4.2 (0.7-26.3)	1.5 (0.6-3.6)	1/1	56.4 (0.277.5)	
826/827: Textile processing occupations	499/1948	25/46	1.8 (1.0-3.2)*	1.6 (1.0-2.6)	15/29	1.5 (0.8-2.9)	1.4 (0.8-2.4)	10/17	2.8 (1.1-6.8)*	1.8 (0.9-3.6)
8278: Occupations in labouring and other elemental work, textile processing	519/1991	5/3	7.8 (1.5-41.2)*	1.8 (0.8-4.2)	4/3	6.2 (1.1-35.2)*	1.7 (0.7-4.0)	1/0		
8551: Patternmaking, marking and cutting occupations	507/1959	17/35	1.6 (0.9-3.0)	1.4 (0.8-2.4)	6/24	0.9 (0.3-2.2)	1.0 (0.5-1.9)	11/11	3.3 (1.3-8.6)*	2.0 (1.0-3.9)*
8595: Painting and decorating occupations, except construction	514/1975	10/19	1.7 (0.8-3.9)	1.4 (0.8-2.7)	8/9	2.9 (1.0-8.1)*	1.8 (0.9-3.6)	2/10	0.6 (0.1-3.2)	1.0 (0.4-2.2)
873: Electrical power, lighting and wire communications equipment erecting, installing and repairing occupations	498/1924	26/70	1.4 (0.9-2.3)	1.4 (0.9-2.1)	12/28	2.2 (1.0-4.7)*	1.7 (0.9-3.1)	14/42	1.1 (0.6-2.1)	1.1 (0.6-1.9)
8799: Other construction trades occupations, n.e.c.	505/1948	19/46	2.0 (1.0-3.6)*	1.7 (1.0-2.8)	12/26	1.9 (0.9-4.1)	1.6 (0.8-2.9)	7/20	2.0 (0.7-5.6)	1.5 (0.7-3.0)
9171: Bus drivers	507/1954	17/40	1.6 (0.9-3.0)	1.5 (0.9-2.5)	9/14	2.9 (1.1-7.3)*	1.9 (0.9-3.6)	8/26	1.1 (0.5-2.5)	1.1 (0.6-2.1)
9171110: Bus driver (motor trans.)	507/1954	17/40	1.6 (0.9-3.0)	1.5 (0.9-2.5)	9/14	2.9 (1.1-7.3)*	1.9 (0.9-3.6)	8/26	1.1 (0.5-2.5)	1.1 (0.6-2.1)
9179118: Dispatcher, motor vehicles (motor trans.)	519/1991	5/3	12.7 (1.6.5)*	1.7 (0.7-4.2)	5/2	13.1 (1.5.1)*	1.7 (0.7-4.2)	0/1		
9310: Foremen/women, material handling and related occupations, n.e.c.	519/1989	5/5	5.1 (1.2-21.5)*	1.8 (0.8-4.0)	2/3	3.6 (0.5-28.4)	1.4 (0.6-3.4)	3/2	8.5 (0.9-81.2)	1.6 (0.6-3.8)
955: Electronic and related communications equipment operating occupations, n.e.c.	515/1985	9/9	3.5 (1.4-9.2)*	2.0 (1.0-4.0)*	7/6	3.8 (1.2-11.6)*	1.9 (0.9-4.0)	2/3	2.8 (0.5-17.4)	1.4 (0.6-3.3)

§ High grade defined by Gleason score ≥8 or 4+3. Occupations selected had at least one significantly elevated or reduced association with prostate cancer among duration of employment categories (ever, <10 years or ≥10 years). Odds ratios estimated using multinomial models adjusted for age, first-degree family history of prostate cancer, ancestry, screening for prostate cancer, annual household income, highest level of education attained, level of physical activity, alcohol intake (drink-years) and body mass index. * The 95% confidence interval excludes the null value. n.e.c.: not elsewhere classified

Table IV. Selected associations between selected industries and risk of prostate cancer (all cancers)

	Never	Ever employed			<10 years employed			≥10 years employed		
SIC code and description	Ca/Co	Ca/Co	OR (95%CI)	OR SB (95%CI)	Ca/Co	OR (95%CI)	OR SB (95%CI)	Ca/Co	OR (95%CI)	OR SB (95%CI)
01: Agricultural industries	1864/1886	73/108	0.6 (0.4-0.9)*	0.6 (0.5-0.9)*	43/66	0.6 (0.4-0.9)*	0.7 (0.4-1.0)*	30/42	0.6 (0.4-1.1)	0.7 (0.4-1.1)
0119: Livestock combination farms	1932/1980	5/14	0.3 (0.1-0.9)*	0.6 (0.3-1.3)	2/7	0.2 (0.0-1.0)	0.7 (0.3-1.6)	3/7	0.4 (0.1-2.4)	0.8 (0.3-1.9)
04: Logging industry	1881/1962	56/32	1.7 (1.0-2.8)*	1.5 (1.0-2.4)	43/23	1.7 (0.9-2.9)	1.5 (0.9-2.4)	13/9	1.8 (0.7-4.7)	1.4 (0.7-2.7)
041: Logging industry	1881/1962	56/32	1.7 (1.0-2.8)*	1.5 (1.0-2.4)	43/23	1.7 (0.9-2.9)	1.5 (0.9-2.4)	13/9	1.8 (0.7-4.7)	1.4 (0.7-2.7)
103: Fruit and vegetable industries	1936/1985	1/9	0.1 (0.0-0.5)*	0.6 (0.3-1.5)	1/7	0.1 (0.0-0.8)*	0.7 (0.3-1.6)	0/2		
1031: Canned and preserved fruit and vegetable industry	1936/1985	1/9	0.1 (0.0-0.5)*	0.6 (0.3-1.5)	1/7	0.1 (0.0-0.8)*	0.7 (0.3-1.6)	0/2		
2619: Other household furniture industries	1922/1989	15/5	3.9 (1.1-13.4)*	1.7 (0.8-3.7)	11/5	3.1 (0.8-11.0)	1.5 (0.7-3.4)	4/0		
27: Paper and allied products industries	1874/1943	63/51	1.3 (0.9-2.0)	1.3 (0.9-1.9)	40/37	1.0 (0.6-1.7)	1.0 (0.6-1.6)	23/14	2.3 (1.1-5.0)*	1.7 (0.9-3.1)
2711: Pulp industry	1911/1972	26/22	1.3 (0.7-2.4)	1.2 (0.7-2.0)	16/18	0.8 (0.4-1.7)	0.9 (0.5-1.6)	10/4	5.6 (1.1-27.5)*	1.6 (0.7-3.8)
3039: Other ornamental and architectural metal products industries	1927/1991	10/3	5.9 (1.0-33.6)*	1.6 (0.7-3.7)	7/2	13.2 (1.38)*	1.5 (0.6-3.7)	3/1	1.6 (0.2-16.1)	
3199: Other machinery and equipment industries n.e.c.	1920/1966	17/28	0.6 (0.3-1.1)	0.7 (0.4-1.2)	11/22	0.4 (0.2-1.0)*	0.6 (0.3-1.1)	6/6	1.0 (0.3-3.6)	1.0 (0.5-2.2)
32: Transportation equipment industries	1809/1831	128/163	0.8 (0.6-1.0)	0.8 (0.6-1.0)	80/90	0.9 (0.6-1.2)	0.9 (0.6-1.2)	48/73	0.7 (0.4-1.0)*	0.7 (0.5-1.0)
321: Aircraft and aircraft parts industry	1881/1897	56/97	0.6 (0.4-0.9)*	0.6 (0.5-0.9)*	32/48	0.7 (0.4-1.2)	0.8 (0.5-1.2)	24/49	0.5 (0.3-0.8)*	0.6 (0.4-0.9)*
3211: Aircraft and aircraft parts industry	1881/1897	56/97	0.6 (0.4-0.9)*	0.6 (0.5-0.9)*	32/48	0.7 (0.4-1.2)	0.8 (0.5-1.2)	24/49	0.5 (0.3-0.8)*	0.6 (0.4-0.9)*
354: Concrete products industries	1925/1987	12/7	4.4 (1.0-18.5)*	1.6 (0.7-3.7)	7/6	2.3 (0.5-11.2)	1.3 (0.6-3.0)	5/1	28.1 (0.89.3)	1.3 (0.5-3.3)
356: Glass and glass products industries	1925/1976	12/18	0.6 (0.3-1.4)	0.8 (0.4-1.4)	9/9	1.2 (0.4-3.4)	1.1 (0.5-2.2)	3/9	0.2 (0.1-1.0)*	0.6 (0.3-1.4)
3561: Primary glass and glass containers industry	1929/1977	8/17	0.5 (0.2-1.2)	0.7 (0.3-1.3)	6/9	0.9 (0.3-2.8)	1.0 (0.4-2.0)	2/8	0.2 (0.0-0.9)*	0.6 (0.3-1.5)
424: Plumbing, heating and air conditioning, mechanical work	1887/1957	50/37	1.7 (1.0-2.8)*	1.5 (1.0-2.4)	23/16	1.7 (0.8-3.6)	1.4 (0.8-2.6)	27/21	1.7 (0.8-3.3)	1.4 (0.8-2.5)
427: Interior and finishing work	1891/1958	46/36	1.7 (1.0-2.9)*	1.5 (1.0-2.4)	17/13	1.8 (0.8-4.4)	1.4 (0.7-2.7)	29/23	1.6 (0.9-3.2)	1.4 (0.8-2.4)
451: Air transport industries	1923/1957	14/37	0.4 (0.2-0.9)*	0.6 (0.3-1.0)	4/10	0.5 (0.1-1.9)	0.8 (0.4-1.8)	10/27	0.4 (0.2-0.9)*	0.6 (0.3-1.1)
4511: Scheduled air transport industry	1923/1959	14/35	0.5 (0.2-0.9)*	0.6 (0.3-1.0)	4/9	0.6 (0.1-2.4)	0.9 (0.4-1.9)	10/26	0.4 (0.2-0.9)*	0.6 (0.3-1.1)
454: Water transport industries	1925/1964	12/30	0.4 (0.2-0.8)*	0.6 (0.3-1.0)	6/21	0.3 (0.1-0.8)*	0.6 (0.3-1.1)	6/9	0.6 (0.2-2.0)	0.8 (0.4-1.8)
4541: Freight and passenger water transport industry	1925/1966	12/28	0.4 (0.2-0.9)*	0.6 (0.3-1.0)	6/19	0.3 (0.1-0.8)*	0.6 (0.3-1.1)	6/9	0.6 (0.2-2.0)	0.8 (0.4-1.8)
4571: Urban transit systems industry	1888/1949	49/45	1.1 (0.7-1.8)	1.1 (0.7-1.6)	21/11	2.4 (1.0-5.8)*	1.7 (0.9-3.2)	28/34	0.8 (0.5-1.4)	0.8 (0.5-1.4)
5512: Trucks and buses, wholesale	1932/1986	5/8	0.5 (0.1-1.8)	0.8 (0.4-1.7)	1/7	0.1 (0.0-1.0)*	0.7 (0.3-1.8)	4/1	2.5 (0.3-23.3)	1.2 (0.5-2.9)
561: Metal and metal products, wholesale	1933/1981	4/13	0.3 (0.1-1.1)	0.6 (0.3-1.4)	1/9	0.1 (0.0-0.8)*	0.7 (0.3-1.7)	3/4	0.9 (0.2-4.5)	1.0 (0.4-2.3)
599: Other products n.e.c., wholesale	1919/1985	18/9	2.5 (1.0-6.0)	1.6 (0.8-3.2)	11/4	4.9 (1.1-21.6)*	1.6 (0.7-3.7)	7/5	1.4 (0.4-4.6)	1.1 (0.5-2.5)
5999: Other products n.e.c., wholesale	1922/1988	15/6	3.4 (1.1-10.4)*	1.7 (0.8-3.6)	8/3	6.5 (0.9-46.1)	1.5 (0.6-3.6)	7/3	2.2 (0.5-8.9)	1.3 (0.6-2.9)

	Never	Ever employed			<10 years employed			≥10 years employed		
SIC code and description	Ca/Co	Ca/Co	OR (95%CI)	OR SB (95%CI)	Ca/Co	OR (95%CI)	OR SB (95%CI)	Ca/Co	OR (95%CI)	OR SB (95%CI)
611: Shoe stores	1931/1980	6/14	0.3 (0.1-1.0)*	0.6 (0.3-1.3)	4/8	0.3 (0.1-1.2)	0.7 (0.3-1.5)	2/6	0.4 (0.1-2.1)	0.8 (0.3-1.9)
6111: Shoe stores	1931/1980	6/14	0.3 (0.1-1.0)*	0.6 (0.3-1.3)	4/8	0.3 (0.1-1.2)	0.7 (0.3-1.5)	2/6	0.4 (0.1-2.1)	0.8 (0.3-1.9)
70: Deposit accepting intermediary industries	1834/1912	103/82	1.3 (0.9-1.9)	1.3 (0.9-1.8)	54/41	1.7 (1.0-2.7)*	1.5 (1.0-2.4)	49/41	1.1 (0.7-1.7)	1.1 (0.7-1.6)
702: Chartered banks and other banking-type intermediaries	1840/1919	97/75	1.4 (1.0-2.0)	1.4 (1.0-1.9)	53/39	1.8 (1.1-2.9)*	1.6 (1.0-2.4)*	44/36	1.1 (0.7-1.8)	1.1 (0.7-1.7)
7021: Chartered banks	1843/1921	94/73	1.4 (1.0-2.0)	1.3 (1.0-1.9)	53/38	1.8 (1.1-2.9)*	1.6 (1.0-2.5)*	41/35	1.1 (0.7-1.8)	1.1 (0.7-1.6)
72: Investment intermediary industries	1916/1977	21/17	1.9 (0.9-4.1)	1.5 (0.8-2.7)	14/7	3.7 (1.1-12.2)*	1.7 (0.8-3.6)	7/10	1.0 (0.3-2.9)	1.0 (0.5-2.1)
8153: Taxation administration, federal	1922/1986	15/8	3.2 (1.0-9.5)*	1.7 (0.8-3.5)	7/4	4.9 (0.7-33.8)	1.4 (0.6-3.4)	8/4	2.4 (0.6-9.4)	1.4 (0.6-3.1)
82: Provincial and territorial government service industries	1781/1892	156/102	1.5 (1.1-2.0)*	1.5 (1.1-2.0)*	72/54	1.3 (0.9-2.0)	1.3 (0.9-1.8)	84/48	1.8 (1.2-2.7)*	1.6 (1.1-2.4)*
822: Protective services (provincial)	1893/1971	44/23	1.8 (1.0-3.2)*	1.6 (1.0-2.5)	23/14	1.6 (0.7-3.4)	1.4 (0.7-2.5)	21/9	2.0 (0.9-4.7)	1.5 (0.8-2.9)
8225: Regulatory services, provincial	1920/1989	17/5	6.1 (1.5-24.8)*	1.8 (0.8-4.1)	8/3	17.0 (1.2.8)*	1.4 (0.6-3.6)	9/2	3.7 (0.8-17.8)	1.5 (0.6-3.4)
825: General administrative services (provincial)	1891/1966	46/28	1.7 (1.0-2.8)	1.5 (0.9-2.4)	21/16	1.0 (0.5-2.1)	1.0 (0.6-1.8)	25/12	3.0 (1.3-7.0)*	1.9 (1.0-3.6)
92: Food and beverage service industries	1786/1796	151/198	0.7 (0.6-0.9)*	0.8 (0.6-1.0)*	91/115	0.7 (0.5-1.0)	0.8 (0.6-1.0)	60/83	0.8 (0.5-1.1)	0.8 (0.6-1.1)
921: Food services	1818/1832	119/162	0.7 (0.6-1.0)*	0.8 (0.6-1.0)*	75/99	0.7 (0.5-1.0)	0.7 (0.5-1.0)	44/63	0.8 (0.5-1.2)	0.8 (0.5-1.2)
9212: Restaurants, unlicensed (including drive-ins)	1906/1931	31/63	0.5 (0.3-0.8)*	0.6 (0.4-0.9)*	24/41	0.6 (0.3-1.0)	0.7 (0.4-1.1)	7/22	0.3 (0.1-0.9)*	0.6 (0.3-1.2)
984: Labour organizations	1921/1974	16/20	0.7 (0.3-1.4)	0.8 (0.4-1.4)	5/13	0.3 (0.1-0.9)*	0.6 (0.3-1.2)	11/7	1.3 (0.5-3.7)	1.2 (0.6-2.4)
9841: Labour organizations	1921/1974	16/20	0.7 (0.3-1.4)	0.8 (0.4-1.4)	5/13	0.3 (0.1-0.9)*	0.6 (0.3-1.2)	11/7	1.3 (0.5-3.7)	1.2 (0.6-2.4)
986: Civic and fraternal organizations	1928/1972	9/22	0.4 (0.2-0.8)*	0.6 (0.3-1.1)	7/13	0.4 (0.2-1.2)	0.7 (0.3-1.4)	2/9	0.2 (0.0-1.2)	0.7 (0.3-1.6)
9861: Civic and fraternal organizations	1928/1972	9/22	0.4 (0.2-0.8)*	0.6 (0.3-1.1)	7/13	0.4 (0.2-1.2)	0.7 (0.3-1.4)	2/9	0.2 (0.0-1.2)	0.7 (0.3-1.6)

§ Industries selected had at least one significantly elevated or reduced association with prostate cancer among duration of employment categories (ever, <10 years or ≥10 years). Odds ratios adjusted for age, first-degree family history of prostate cancer, ancestry, screening for prostate cancer, annual household income, highest level of education attained, level of physical activity, alcohol intake and body mass index. * The 95% confidence interval excludes the null value. n.e.c.: not elsewhere classified

Table V. Selected associations between selected industries and risk of high grade prostate cancer

	Never	Ever employed			<10 years employed			≥10 years employed		
SIC code and description	**Ca/Co**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**
0111: Dairy farms	518/1988	6/6	3.5 (0.9-12.8)	1.7 (0.8-3.7)	6/4	4.5 (1.1-18.9)*	1.8 (0.8-4.0)	0/2		
016: Horticultural specialties	520/1988	4/6	2.7 (0.7-11.3)	1.5 (0.7-3.4)	4/2	9.8 (1.1-87.4)*	1.6 (0.7-4.0)	0/4		
04: Logging industry	503/1962	21/32	2.1 (1.1-3.9)*	1.8 (1.1-3.0)*	14/23	1.8 (0.9-3.7)	1.5 (0.9-2.8)	7/9	3.1 (1.0-9.3)*	1.8 (0.9-3.7)
041: Logging industry	503/1962	21/32	2.1 (1.1-3.9)*	1.8 (1.1-3.0)*	14/23	1.8 (0.9-3.7)	1.5 (0.9-2.8)	7/9	3.1 (1.0-9.3)*	1.8 (0.9-3.7)
0412: Contract logging industry	518/1986	6/8	3.7 (1.1-12.5)*	1.8 (0.8-3.9)	4/6	4.2 (0.9-19.0)	1.7 (0.8-3.8)	2/2	2.7 (0.3-21.3)	
104: Dairy products industries	512/1980	12/14	2.5 (1.0-6.0)*	1.7 (0.9-3.3)	7/11	1.7 (0.6-5.2)	1.4 (0.6-2.8)	5/3	5.3 (1.0-28.0)*	1.7 (0.7-4.0)
25: Wood industries	504/1955	20/39	1.9 (1.0-3.6)*	1.7 (1.0-2.8)	17/31	2.2 (1.1-4.4)*	1.8 (1.0-3.1)*	3/8	1.0 (0.2-4.3)	1.1 (0.5-2.5)
2619: Other household furniture industries	518/1989	6/5	5.5 (1.4-21.9)*	1.9 (0.9-4.3)	3/5	3.0 (0.6-14.7)	1.5 (0.6-3.4)	3/0		
291: Primary steel industries	511/1969	13/25	2.1 (1.0-4.5)*	1.7 (0.9-3.0)	6/17	1.5 (0.6-4.2)	1.3 (0.6-2.7)	7/8	3.2 (1.1-9.8)*	1.8 (0.9-3.8)
2919: Other primary steel industries	518/1988	6/6	4.5 (1.2-16.3)*	1.9 (0.8-4.1)	2/4	3.2 (0.4-23.0)	1.4 (0.6-3.4)	4/2	5.7 (1.0-32.7)	1.7 (0.7-4.0)
3039: Other ornamental and architectural metal products industries	520/1991	4/3	8.3 (1.2-55.9)*	1.7 (0.7-4.1)	4/2	27.1 (2.3.0)*	1.8 (0.7-4.4)	0/1		
3331: Lighting fixture industry	519/1989	5/5	5.2 (1.2-21.9)*	1.8 (0.8-4.1)	3/4	4.8 (0.8-29.3)	1.6 (0.7-3.8)	2/1	5.8 (0.5-66.3)	
4123: Hydroelectric power plants and related structures(except transmission lines)	518/1987	6/7	4.9 (1.3-17.7)*	1.9 (0.9-4.2)	5/6	4.3 (1.1-16.4)*	1.8 (0.8-4.0)	1/1	26.7 (0.19.5)	
4214: Excavating and grading	516/1980	8/14	2.1 (0.8-5.3)	1.5 (0.8-3.0)	5/5	5.9 (1.2-29.5)*	1.8 (0.8-4.1)	3/9	1.0 (0.3-4.0)	1.1 (0.5-2.4)
4241: Plumbing	512/1966	12/28	1.8 (0.9-3.9)	1.5 (0.8-2.8)	8/11	3.2 (1.1-8.9)*	1.9 (0.9-3.8)	4/17	1.0 (0.3-3.2)	1.1 (0.5-2.3)
456: Truck transport industries	492/1917	32/77	1.5 (1.0-2.4)	1.4 (0.9-2.2)	7/35	0.8 (0.4-2.0)	1.0 (0.5-1.8)	25/42	2.0 (1.1-3.5)*	1.7 (1.1-2.8)*
4571: Urban transit systems industry	509/1949	15/45	1.2 (0.7-2.3)	1.2 (0.7-2.0)	8/11	3.5 (1.2-9.8)*	1.9 (0.9-4.0)	7/34	0.7 (0.3-1.7)	0.9 (0.5-1.7)
492: Gas distribution systems industry	519/1988	5/6	3.7 (1.0-13.8)*	1.7 (0.8-3.8)	1/2	2.0 (0.2-26.0)		4/4	4.6 (1.0-21.3)	1.7 (0.8-3.9)
4921: Gas distribution systems industry	519/1988	5/6	3.7 (1.0-13.8)*	1.7 (0.8-3.8)	1/2	2.0 (0.2-26.0)		4/4	4.6 (1.0-21.3)	1.7 (0.8-3.9)
599: Other products n.e.c., wholesale	518/1985	6/9	3.1 (1.0-9.5)*	1.7 (0.8-3.6)	2/4	3.6 (0.5-25.6)	1.4 (0.6-3.5)	4/5	2.8 (0.7-10.9)	1.5 (0.7-3.4)
5999: Other products n.e.c., wholesale	518/1988	6/6	5.0 (1.4-18.4)*	1.9 (0.9-4.2)	2/3	6.4 (0.6-64.8)	1.5 (0.6-3.7)	4/3	4.3 (0.9-21.0)	1.7 (0.7-3.8)
6212: Household furniture stores (without appliances and furnishings)	516/1986	8/8	3.1 (1.1-8.8)*	1.8 (0.9-3.7)	5/5	3.2 (0.8-12.3)	1.6 (0.7-3.6)	3/3	2.9 (0.6-14.7)	1.5 (0.6-3.4)
633: Gasoline service stations	506/1968	18/26	2.8 (1.4-5.5)*	2.1 (1.2-3.6)*	16/22	2.9 (1.4-6.0)*	2.1 (1.2-3.7)*	2/4	2.0 (0.3-11.6)	1.3 (0.6-3.1)
6331: Gasoline service stations	506/1968	18/26	2.8 (1.4-5.5)*	2.1 (1.2-3.6)*	16/22	2.9 (1.4-6.0)*	2.1 (1.2-3.7)*	2/4	2.0 (0.3-11.6)	1.3 (0.6-3.1)
70: Deposit accepting intermediary industries	497/1912	27/82	1.4 (0.9-2.3)	1.4 (0.9-2.1)	16/41	1.9 (1.0-3.7)*	1.6 (1.0-2.8)	11/41	1.1 (0.5-2.1)	1.1 (0.6-1.9)
702: Chartered banks and other banking-type intermediaries	498/1919	26/75	1.5 (0.9-2.5)	1.4 (0.9-2.2)	16/39	2.0 (1.1-3.9)*	1.7 (1.0-2.9)	10/36	1.1 (0.5-2.3)	1.1 (0.6-2.0)
7021: Chartered banks	498/1921	26/73	1.6 (1.0-2.6)	1.5 (0.9-2.3)	16/38	2.1 (1.1-4.0)*	1.7 (1.0-3.0)*	10/35	1.2 (0.6-2.4)	1.1 (0.6-2.1)
72: Investment intermediary industries	517/1977	7/17	2.8 (1.0-7.5)*	1.8 (0.9-3.5)	4/7	4.5 (1.0-19.4)*	1.7 (0.8-3.9)	3/10	2.0 (0.5-8.0)	1.4 (0.6-3.0)
729: Other investment intermediaries	520/1989	4/5	6.0 (1.3-27.7)*	1.8 (0.8-4.2)	2/0			2/5	3.0 (0.5-18.3)	1.4 (0.6-3.4)

	Never	Ever employed			<10 years employed			≥10 years employed		
SIC code and description	**Ca/Co**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**	**Ca/Co**	**OR (95%CI)**	**OR SB (95%CI)**
7299: Other investment intermediaries n.e.c.	520/1990	4/4	6.1 (1.3-27.9)*	1.8 (0.8-4.2)	2/0			2/4	3.0 (0.5-18.3)	1.4 (0.6-3.4)
81: Federal government service industries	469/1734	55/260	0.8 (0.6-1.1)	0.8 (0.6-1.1)	33/183	0.7 (0.4-1.0)*	0.7 (0.5-1.0)	22/77	1.2 (0.7-1.9)	1.2 (0.7-1.8)
8152: Finance and economic administration, federal	520/1987	4/7	7.7 (1.5-38.7)*	1.9 (0.8-4.4)	1/5	3.1 (0.2-42.6)	1.3 (0.5-3.2)	3/2	14.4 (1.6.0)*	1.7 (0.7-4.3)
8164: Recreation and culture administration, federal	521/1993	3/1	122.3 (0.554.8)		3/1	137.2 (63.6.9)*		0/0		
82: Provincial and territorial government service industries	482/1892	42/102	1.7 (1.2-2.6)*	1.6 (1.1-2.4)*	22/54	1.6 (0.9-2.8)	1.5 (0.9-2.4)	20/48	1.8 (1.0-3.2)*	1.6 (1.0-2.7)*
8259: Other general administrative services, provincial	519/1992	5/2	8.0 (1.5-42.9)*	1.8 (0.8-4.3)	3/1	8.9 (0.9-88.6)		2/1	7.5 (0.6-86.4)	
83: Local government service industries	483/1856	41/138	1.2 (0.8-1.7)	1.1 (0.8-1.6)	20/44	1.9 (1.1-3.4)*	1.7 (1.0-2.8)*	21/94	0.8 (0.5-1.4)	0.9 (0.6-1.4)
8362: Social service administration, municipal	519/1988	5/6	4.0 (1.1-15.2)*	1.8 (0.8-3.9)	2/3	6.6 (0.7-62.7)	1.5 (0.6-3.7)	3/3	2.9 (0.6-14.8)	1.5 (0.6-3.4)
855: Museums and archives	519/1990	5/4	3.4 (0.8-13.7)	1.6 (0.7-3.6)	5/3	4.9 (1.1-22.4)*	1.8 (0.8-4.0)	0/1		
8551: Museums and archives	519/1990	5/4	3.4 (0.8-13.7)	1.6 (0.7-3.6)	5/3	4.9 (1.1-22.4)*	1.8 (0.8-4.0)	0/1		
864: Non-institutional social services	516/1975	8/19	1.7 (0.7-4.3)	1.4 (0.7-2.8)	6/11	3.4 (1.1-10.6)*	1.8 (0.9-3.8)	2/8	0.6 (0.1-3.0)	1.0 (0.4-2.2)
8642: Child welfare services	519/1988	5/6	2.9 (0.8-10.7)	1.6 (0.7-3.5)	3/1	62.2 (1.83.3)*		2/5	1.0 (0.2-5.6)	1.1 (0.5-2.6)
9212: Restaurants, unlicensed (including drive-ins)	516/1931	8/63	0.4 (0.2-1.0)*	0.6 (0.3-1.2)	6/41	0.5 (0.2-1.2)	0.7 (0.4-1.4)	2/22	0.3 (0.1-1.4)	0.8 (0.3-1.8)
961: Motion picture, audio and video production and distribution	518/1985	6/9	2.8 (0.9-8.6)	1.7 (0.8-3.5)	1/3	1.2 (0.1-11.7)		5/6	3.9 (1.0-14.8)*	1.8 (0.8-3.9)

§ High grade defined by Gleason score ≥8 or 4+3. Industries selected had at least one significantly elevated or reduced association with prostate cancer among duration of employment categories (ever, <10 years or ≥10 years). Odds ratios estimated using multinomial models adjusted for age, first-degree family history of prostate cancer, ancestry, screening for prostate cancer, annual household income, highest level of education attained, level of physical activity, alcohol intake and body mass index. * The 95% confidence interval excludes the null value. n.e.c.: not elsewhere classified

A1.9. References

Alavanja MCR, Blair A, Merkle S *et al.* (1989) Mortality among Forest and Soil Conservationists. Arch Environ Health; 44: 94-101.

Alavanja MCR, Bonner MR. (2012) Occupational Pesticide Exposures and Cancer Risk: A Review. J Toxicol Environ Health, Part B; 15: 238-63.

Aronson KJ, Siemiatycki J, Dewar R *et al.* (1996) Occupational Risk Factors for Prostate Cancer: Results from a Case-Control Study in Montréal, Québec, Canada. Am J Epidemiol; 143: 363-73.

Band PR, Abanto Z, Bert J *et al.* (2011) Prostate cancer risk and exposure to pesticides in British Columbia Farmers. The Prostate; 71: 168-83.

Band PR, Le ND, Fang R *et al.* (1999) Identification of occupational cancer risks in British Columbia. Part II: A population-based case-control study of 1516 Prostatic cancer cases. J Occup Environ Med; 41: 233-47.

Blanc-Lapierre A, Spence A, Karakiewicz PI *et al.* (2015) Metabolic syndrome and prostate cancer risk in a population-based case–control study in Montreal, Canada. BMC Public Health; 18: 913.

Buja A, Lange JH, Perissinotto E *et al.* (2005) Cancer incidence among male military and civil pilots and flight attendants: an analysis on published data. Toxicol Ind Health; 21: 273-82.

Canadian Cancer Society's Advisory Committee on Cancer Statistics. (2015) Canadian Cancer Statistics 2015. Toronto, ON: Canadian Cancer Society.

Cogliano VJ, Baan R, Straif K *et al.* (2011a) Preventable Exposures Associated With Human Cancers. J Natl Cancer Inst.

Cogliano VJ, Baan R, Straif K *et al.* (2011b) Preventable Exposures Associated With Human Cancers. J Natl Cancer Inst; 203: 1827-39.

Corbin M, Richiardi L, Vermeulen R *et al.* (2012) Hierarchical Regression for Multiple Comparisons in a Case-Control Study of Occupational Risks for Lung Cancer. PLoS ONE; 7: e38944.

Department of Employment and Immigration. (1971) Canadian classification dictionnary of occupations, vol. 1. Ottawa, ON: Department of Employment and Immigration.

Depczynski J, Lower T. (2014) A review of prostate cancer incidence and mortality studies of farmers and non-farmers, 2002–2013. Cancer Epidemiol; 38: 654-62.

Doolan G, Benke G, Giles G. (2014) An Update on Occupation and Prostate Cancer. Asian Pac J Cancer Prev; 15: 501-16.

Driscoll TR, Carey RN, Peters S *et al.* (2016) The Australian Work Exposures Study: Occupational Exposure to Polycyclic Aromatic Hydrocarbons. Ann Occup Hyg; 60: 124-31.

Dryson E, t Mannetje A, Walls C *et al.* (2008) Case-control study of high risk occupations for bladder cancer in New Zealand. Int J Cancer; 122: 1340-46.

Ferlay J, Soerjomataram I, Dikshit R *et al.* (2015) Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer; 136: E359-E86.

Finkelstein MM. (1998) Cancer incidence among Ontario police officers. Am J Ind Med; 34: 157-62.

Greenland S, Poole C. (1994) Empirical-Bayes and Semi-Bayes Approaches to Occupational and Environmental Hazard Surveillance. Arch Environ Health; 49: 9-16.

IARC. (2013) Non-Ionizing Radiation, Part 2: Radiofrequency Electromagnetic Fields. Vol. 102. Lyon, France: International Agency for Research on Cancer.

Jack RJ, Oliver M. (2008) A Review of Factors Influencing Whole-Body Vibration Injuries in Forestry Mobile Machine Operators. Int J Forest Eng; 19: 51-65.

Järvholm B, Silverman D. (2003) Lung cancer in heavy equipment operators and truck drivers with diesel exhaust exposure in the construction industry. Occup Environ Med; 60: 516-20.

Jones RR, Barone-Adesi F, Koutros S *et al.* (2015) Incidence of solid tumours among pesticide applicators exposed to the organophosphate insecticide diazinon in the Agricultural Health Study: an updated analysis. Occup Environ Med; 72: 496-503.

Kangas J. (2011) Chemical Hazards. In Poschen, P, ed. Chapter 68, Forestry. . Encyclopedia of Occupational Health and Safety, Geneva: International Labor Organization.

Koutros S, Andreotti G, Berndt SI *et al.* (2011) Xenobiotic-metabolizing gene variants, pesticide use, and the risk of prostate cancer. Pharmacogenet Genomics; 21: 615-23.

Koutros S, Beane Freeman LE, Lubin JH *et al.* (2013) Risk of Total and Aggressive Prostate Cancer and Pesticide Use in the Agricultural Health Study. Am J Epidemiol; 177: 59-74.

Krstev S, Baris D, Stewart PA *et al.* (1998) Risk for prostate cancer by occupation and industry: a 24-state death certificate study. Am J Ind Med; 34: 413-20.

Mitler MM, Miller JC, Lipsitz JJ *et al.* (1997) The Sleep of Long-Haul Truck Drivers. N Engl J Med; 337: 755-62.

Nadalin V, Kreiger N, Parent M-E *et al.* (2012) Prostate Cancer and Occupational Whole-Body Vibration Exposure. Ann Occup Hyg; 56: 968-74.

Parent M-É, Siemiatycki J. (2001) Occupation and Prostate Cancer. Epidemiol Rev; 23: 138-43.

Patel A, Klein E. (2009) Risk factors for prostate cancer. Nat Clin Pract Urol; 6: 87-95.

Peters CE, Demers PA, Kalia S *et al.* (2016) Occupational exposure to solar ultraviolet radiation and the risk of prostate cancer.

Potti A, Panwalkar AW, Langness E. (2003) Prevalence of pesticide exposure in young males (</= 50 years) with adenocarcinoma of the prostate. J Carcinog; 2: 4.

Pukkala E, Martinsen JI, Lynge E *et al.* (2009) Occupation and cancer – follow-up of 15 million people in five Nordic countries. Acta Oncol; 48: 646-790.

Pukkala E, Martinsen JI, Weiderpass E *et al.* (2014) Cancer incidence among firefighters: 45 years of follow-up in five Nordic countries. Occup Environ Med; 71: 398-404.

Ragin C, Davis-Reyes B, Tadesse H *et al.* (2013) Farming, Reported Pesticide Use, and Prostate Cancer. Am J Mens Health; 7: 102-09.

Ramanakumar AV, Nadon L, Siemiatycki J. (2008) Exposures in painting related occupations and risk of selected cancers: Results from a case–control study in montreal. Am J Ind Med; 51: 419-27.

Rao D, Yu H, Bai Y *et al.* (2015) Does night-shift work increase the risk of prostate cancer? a systematic review and meta-analysis. Onco Targets Ther; 8: 2817-26.

Rybicki BA, Neslund-Dudas C, Nock NL *et al.* (2006) Prostate cancer risk from occupational exposure to polycyclic aromatic hydrocarbons interacting with the GSTP1 Ile105Val polymorphism. Cancer Detect Prev; 30: 412-22.

Schwartz E. (1987) Proportionate mortality ratio analysis of automobile mechanics and gasoline service station workers in New Hampshire. Am J Ind Med; 12: 91-99.

Seidler A, Heiskel H, Bickeboller R *et al.* (1998) Association between diesel exposure at work and prostate cancer. Scand J Work Environ Health; 24: 486-94.

Sharma-Wagner S, Chokkalingam AP, Malker HS *et al.* (2000) Occupation and prostate cancer risk in Sweden. J Occup Environ Med; 42: 517-25.

Sharpe CR, Siemiatycki J, Parent M-É. (2001) Activities and Exposures during Leisure and Prostate Cancer Risk. Cancer Epidemiol Biomarkers Prev; 10: 855-60.

Siegel RL, Miller KD, Jemal A. (2016) Cancer statistics, 2016. CA Cancer J Clin; 66: 7-30.

Siemiatycki J, Day NE, Fabry J *et al.* (1981) Discovering Carcinogens in the Occupational Environment: A Novel Epidemiologic Approach. J Natl Cancer Inst; 66: 217-25.

Siemiatycki J, Richardson L, Straif K *et al.* (2004) Listing occupational carcinogens. Environmental Health Perspectives; 112: 1447-59.

Sritharan J, Demers PA, Harris SA *et al.* (2016) Natural resource-based industries and prostate cancer risk in Northeastern Ontario: a case–control study. Occup Environ Med; 73: 506-11.

Srogi K. (2007) Monitoring of environmental exposure to polycyclic aromatic hydrocarbons: a review. Environ Chem Lett; 5: 169-95.

Statistics Canada. (1980) Standard Industrial Classification, 1980. Ottawa, ON: Statistics Canada, Standards Division.

Su LJ, Fontham ETH. (2014) Prostate cancer: an occupational hazard? Future Oncol; 10: 903-06.

Su LJ, Sun W, Fontham ET *et al.* (2013) Occupational risk factors for high aggressive prostate cancer in a population-based study of African Americans and European Americans [abstract]. Cancer Prev Res; 6 (Suppl 11): B63.

Teschke K, Olshan AF, Daniels JL *et al.* (2002) Occupational exposure assessment in case-control studies: opportunities for improvement. Occupational and Environmental Medicine; 59: 575-94.

van der Gulden JW, Kolk JJ, Verbeek AL. (1995) Work environment and prostate cancer risk. Prostate; 27: 250-57.

Wirth M, Vena JE, Smith EK *et al.* (2013) The epidemiology of cancer among police officers. Am J Ind Med; 56: 439-53.

Wright JL, Salinas CA, Lin DW *et al.* (2009) Prostate Cancer Specific Mortality and Gleason 7 Disease Differences in Prostate Cancer Outcomes Between Cases With Gleason 4 + 3 and Gleason 3 + 4 Tumors in a Population Based Cohort. J Urol; 182: 2702-07.

Young E, Kreiger N, Purdham J *et al.* (2009) Prostate cancer and driving occupations: could whole body vibration play a role? Int Arch Occup Environ Health; 82: 551-56.

Zeegers MP, Friesema IH, Goldbohm RA *et al.* (2004) A prospective study of occupation and prostate cancer risk. J Occup Environ Med; 46: 271-79.

Annexe 2. Classifications professionnelles et industrielles utilisées dans CANJEM

Classifications professionnelles et niveaux hiérarchiques	Nombre de groupes
Classification Internationale Type des Professions, 1968 (Internationale)[1]	
Grands groupes	8
Sous-groupes	81
Groupes de base	282
Professions	1504
Classification Canadienne Descriptive des Professions, 1971 (Canada)	
Grands groupes	23
Sous-groupes	81
Groupes de base	500
Professions	7907
Classification Nationale des Professions, 2011 (Canada)	
Grandes catégories professionnelles	10
Grands groupes	40
Groupes intermédiaires	140
Groupes de base	500
Standardized Occupational Classification, 2010 (États-Unis)	
Major groups	23
Minor groups	97
Broad occupations	461
Detailed occupations	840
Classifications industrielles et niveaux hiérarchiques	**Nombre de groupes**
Classification internationale type, par industrie, de toutes les branches d'activité économique (Internationale) révision 2, 1968[2]	
Branches	9
Catégories	33
Classes	71
Groupes	159
Classification type des industries, 1980 (Canada)	
Division	18
Grands groupes	76
Groupes	318
Industries	860
Système de classification des industries de l'Amérique du Nord, 2012 (Canada/États-Unis)	

Secteurs	20
Sous-secteurs	102
Groupes	323
Classes	711
Classes Canadiennes	922

1. Inclut les membres des forces armées comme catégorie distincte pour chacun des niveaux
2. Exclut les catégories correspondant à la branche « Activités mal désignées » (CITI 0)

Annexe 3. Liste des 258 agents inclus dans CANJEM, stratifiés par catégorie et groupe chimique

Agent
<u>**Substances inorganiques**</u>
Poussières inorganiques
Poussières d'abrasif
Poussières d'isolants inorganiques
Poussière métallique
Amiante chrysotile
Amiante amphibole
Silice cristalline
Ciment Portland
Poussière de verre
Fibres de verre
Talc industriel
Poussière de brique
Poussière d'argile
Poussière de béton
Poussière de brique refractaire
Poussière de bronze
Poussière de laiton
Poussière d'acier
Poussière d'acier doux
Pigments inorganiques
Fibres de laine minérale
Matières de charge
Cendres
Mica
Talc cosmétique
Carbonate de sodium
Hydrosulfite de sodium
Alumine
Carbure de Silicium
Soufre
Oxyde de calcium
Sulfate de calcium
Carbonate de calcium
Dioxyde de titane
Oxydes de fer
Oxyde de zinc
Oxydes de Plomb
Carbonate basique de Plomb
Chromate de Plomb
Gaz inorganiques
Hydrogène
Monoxyde de carbone
Ammoniac
Oxydes d'azote
Ozone
Fluorure d'hydrogène
Dioxyde de soufre
Sulfure d'hydrogène
Chlore
Chlorure d'hydrogène
Dioxyde de chlore
Fumées inorganiques
Fumées de soudage au gaz
Fumées de soudage à l'arc
Fumées de brasage tendre
Fumées d'oxydes métalliques
Fumées d'aluminium
Fumées d'oxyde de calcium
Fumées de dioxyde de titane
Fumées de chrome
Fumées de manganèse
Fumées de fer
Fumées de nickel
Fumées de cuivre
Fumées de zinc
Fumées d'argent
Fumées de cadmium
Fumées d'étain
Fumées de plomb
Liquides et vapeurs inorganiques
Acides inorganiques en solution

Agent
Alcalis caustiques en solution
Solution de placage
Acide nitrique
Peroxyde d'hydrogène
Acide phosphorique
Acide sulfurique
Composés métalliques
Composés du chrome (VI)
Composés du béryllium
Composés du magnésium
Composés de l'aluminium
Composés du titane
Composés du vanadium
Composés du chrome
Composés du manganèse
Composés du fer
Composés du cobalt
Composés du nickel
Composés du cuivre
Composés du zinc
Composés de l'arsenic
Composés du sélénium
Composés de l'argent
Composés du cadmium
Composés de l'étain
Composés de l'antimoine
Composés du tungstène
Composés du mercure
Composés du plomb
Autres substances inorganiques
Cyanures
Fluorures
Hypochlorites
Nitrates
<u>Substances organiques</u>
Poussières organiques
Teintures et pigments organiques
Poussière de coton
Fibres de laine
Fibres de soie
Poussière de bois

Agent
Poussière de grain
Poussière de farine
Poussière de fourrure
Fibres de lin
Poussière de liège
Poussière de cheveux
Poussière d'amidon
Poussière de sucre
Poussière de feutre
Poussière de cuir
Poussière de tabac
Caoutchouc naturel
Acide tannique
Fibres synthétiques
Poussières de matière plastique
Fibres de rayonne
Fibres acryliques
Fibres polyester
Fibres de nylon
Fibres d'acétate
Acétate de cellulose
Nitrate de cellulose
Polyéthylène
Polypropylène
Polystyrène
Poly(chlorure de vinyle)
Poly(acétate de vinyle)
Polyamides
Polyacrylates
Époxy
Phénol-formaldéhyde
Urée-formaldéhyde
Mélamine-formaldéhyde
Polyuréthanes
Polyesters
Caoutchouc styrène-butadiène
Polychloroprène
Fibres textiles
Cellulose
Poussière de caoutchouc
Gaz organiques

Agent
Cyanure d'hydrogène
Gaz naturel
Méthane
Propane
Formaldéhyde
Oxyde d'éthylène
Acétylène
Chlorure de vinyle
Phosgène
Gaz anesthésiants
Gaz propulseurs
Alcanes (C1-C4)
Aldéhydes aliphatiques
Hydrocarbures aliphatiques insaturés
Liquides et vapeurs organiques
Méthanol
Éthanol
Éthylène glycol
Isopropanol
Acide acétique
Acide formique
Éther diéthylique
Tétrachlorure de carbone
Chloroforme
Dichlorométhane
1,1,1-Trichloroéthane
Disulfure de carbone
Trichloréthylène
Perchloroéthylène
Méthacrylate de méthyle
Acétone
Benzène
Toluène
Xylène
Styrène
Phénol
Colles animales et végétales
Essence de térébenthine
Huile de lin
Adhésifs synthétiques
Solvants organiques

Agent
Cires, polis
Essence au plomb
Kérosène
Carburant diesel
Mazout
Essences minérales après 1970
Pétrole brut
Huiles et graisses lubrifiantes
Asphalte
Goudron et brai de houille
Fluide hydraulique
Autres huiles minérales
Essence d'aviation
Essences minérales avant 1970
Biphényles polychlorés (BPC)
Fluides de coupe avant 1955
Fluides de coupe après 1955
Alcanes (C18+)
Alcanes (C5-C17)
Alcools aliphatiques
Alcanes chlorés
Alcènes chlorés
Esters aliphatiques
Cétones aliphatiques
Fluorocarbones
Éthers de glycol
Hydrocarbures aromatiques polycycliques (HAP) de toute origine
HAP dérivés d'autres sources
HAP dérivés du bois
HAP dérivés du pétrole
HAP dérivés du charbon
Benzo[a]pyrène
Hydrocarbures aromatiques mononucléaires
Alcools aromatiques
Autres substances organiques
Alkydes
Nitroglycérine
Cyclotriméthylènetrinitramine (RDX)
Trinitrotoluène

Agent
Amines aromatiques
Phtalates
Isocyanates
<u>Radiations, champs électriques et magnétiques</u>
Radiation ionisante
Radiation ionisante
Radiofréquence, micro-ondes
Radiofréquence, micro-ondes
Rayonnement ultraviolet
Rayonnement ultraviolet
<u>Mélanges inorganiques et organiques</u>
Poussières contenant des substances inorganiques et organiques
Poussière de charbon
Noir de carbone
Suie
Poussière de coke
Poussière de graphite
Mélanges de gaz inorganiques et organiques
Gaz de houille
Fumées de combustion inorganiques et organiques
Autres fumées de pyrolyse
Fumées de cuisson
Gaz d'échappement
Produits de combustion du charbon
Gaz d'échappement diesel
Produits de combustion de combustible liquide
Produits de combustion du bois
Produits de combustion du gaz naturel
Gaz d'échappement propane
Fumées de la pyrolyse de plastique
Fumées de la pyrolyse de caoutchouc
Produits de combustion du Propane
Produits de combustion du coke
Mélanges liquides de substances inorganiques et organiques
Autres peintures, vernis
Vernis et teintures

Agent
Encres
Revêtements métalliques
<u>Catégories générales</u>
Agents de nettoyage
Agents de nettoyage
Engrais
Engrais
Pesticides
Dichlorodiphényltrichloroéthane (DDT)
Créosote
Pesticides
Biocides
Biocides
Décolorants
Décolorants