



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Mention : « Sciences de la Vie et de la Santé »

par Gautier MATER

**Caractérisation des mesures d'exposition à des produits
chimiques dans les bases de données françaises COLCHIC et
SCOLA pour la prévention des maladies professionnelles**

13 décembre 2016

Membres du jury :

Rapporteurs : Monsieur Alain BERGERET Professeur, Université de LYON 1
 Monsieur Patrick BROCHARD Professeur, Université de BORDEAUX

Examineurs : Monsieur Alain GARRIGOU Professeur, Université de BORDEAUX
 Monsieur Francis GUILLEMIN Professeur, Université de LORRAINE
 Monsieur Jérôme LAVOUE Professeur, Université de MONTREAL
 Monsieur Christophe PARIS Professeur, Université de RENNES

Membres invités : Monsieur Raymond VINCENT Chercheur, INRS, VANDOEUVRE
 Monsieur Pascal WILD Chercheur, INRS, VANDOEUVRE

EA 7298 – Interactions gènes – Risques environnementaux et effets sur la santé (INGRES)
Faculté de médecine, 9 avenue de la Forêt de Haye
BP 184 -54505 VANDOEUVRE LES NANCY CEDEX

*« Avec de l'entraînement,
je vais acquérir une force extraordinaire dans le bout des
doigts ; quand ma main croche,
rappelle-toi que je ne suis pas près de dévisser. »*

Roger Frison-Roche
Premier de cordée

Ce travail a été réalisé au laboratoire INGRES (Interactions gènes – Risques environnementaux et effets sur la santé), EA 7298 Université de Lorraine, à la Faculté de Médecine, sous la direction du Professeur Christophe Paris et la codirection du Professeur Jérôme Lavoué Département de Santé environnementale et santé au travail Université de Montréal (Québec, Canada).

Avant d'exposer mes travaux présentés dans ce mémoire, je tiens à remercier celles et ceux qui ont permis leur réalisation.

Mes remerciements vont au Professeur Christophe Paris, à qui je tiens à exprimer ma gratitude, pour son accueil au sein du laboratoire INGRES, sa bienveillance, son soutien et ses encouragements tout au long de mes travaux de thèse.

Je tiens à adresser toute ma reconnaissance au Professeur Jérôme Lavoué pour son soutien technique, sa disponibilité, ses encouragements malgré les milliers de kilomètres que sépare Nancy de Montréal. Il m'a guidé de ses conseils permanents et judicieux et m'a communiqué son enthousiasme pour la recherche.

Mes remerciements vont également à Monsieur Bergeret, Professeur à l'Université de Lyon, Monsieur Brochard, Professeur à l'Université de Bordeaux, Monsieur Garrigou, Professeur à l'Université de Bordeaux et Monsieur Guillemin, Professeur à l'Université de Lorraine pour m'avoir fait l'honneur d'accepter de participer à ce jury.

Je remercie l'Institut National de Recherche et Sécurité (INRS) pour m'avoir permis, dans le cadre de mes activités professionnelles, de mener ces travaux de recherche au travers d'une thèse.

Je tiens à remercier M. Benoît Courier, Chef du Département de Métrologie des Polluants à l'INRS, de m'avoir soutenu dans ma démarche, pour les discussions scientifiques très enrichissantes que nous avons eues, pour ta disponibilité, ton soutien et ta bonne humeur.

Merci également à l'ensemble des personnes du Département de Métrologie des Polluants qui m'ont conforté dans mes travaux.

Je tiens par la même occasion à remercier l'ensemble des personnes du laboratoire de Caractérisation du Risque Chimique de l'INRS, « mes collaborateurs », qui m'ont soutenu

tout au long de ces 4 années : Sarah Burzoni, Bertrand Honnert, Brigitte Jeandel, Marilyne L'Huillier et Barbara Savary.

Je remercie également Frédéric Clerc, Carole Invernizzi, Serge Morillon, Raymond Vincent et Pascal Wild sur qui j'ai pu compter pendant mes travaux.

Merci à l'ensemble des personnes du Centre de Recherche du Centre Hospitalier Universitaire de Montréal (CR CHUM), qui m'ont, lors de mes séjours à Montréal, épaulé dans mes travaux : Philippe Sarazin, Jean-François Sauvé, Mawoulé Sant'Anna, Spéro Abattan, Thomas Remen...

Je remercie tous mes amis « extérieurs », avec qui j'ai pu passer des moments de décompression.

Je tiens à remercier très chaleureusement mon épouse Florence et mes enfants Paul, Jules et Jeanne de m'avoir accompagné tout au long de ma thèse et soutenu dans cette belle aventure.

Un immense merci à ma Maman, mon frère Clément, Julie, Justin, Louise, Michèle, Guy, Pierre-André, Camille, Jérôme, Alix, Alban, Etienne et Naïma pour leur accompagnement tout au long de ces 4 années.

Enfin, je dédie ce travail à Papa, parti trop tôt suite à son cancer. Malgré la maladie, Papa avait toujours quelques mots pour me motiver et m'encourager afin que je donne le meilleur de moi-même dans ce travail de thèse.

TABLE DES MATIERES

Abréviation	1
Avant propos	2
Organisation de la thèse.....	5
CHAPITRE I. Introduction générale	4
I.1. Contexte	4
I.2. Les bases de données d'exposition professionnelle.....	4
I.2.1. Les limites de l'utilisation des bases de données nationales comme sources d'information sur l'exposition	8
I.2.2. Origine des biais potentiels.....	8
I.2.3. Biais rapportés dans la littérature	10
CHAPITRE II. Objectifs des travaux de thèse	12
II.1. Objectif général	12
II.2. Objectifs spécifiques	12
CHAPITRE III. Méthodes.....	14
III.1. Présentation des bases de données d'exposition professionnelle françaises	14
III.1.1. COLCHIC	14
III.1.1.1. Généralités.....	14
III.1.1.2. Exploitation des données de COLCHIC	16
III.1.1.3. Les évolutions de COLCHIC.....	16
III.1.2. SCOLA	17
III.1.2.1. Généralités.....	17
III.1.2.2. Exploitation des données de SCOLA.....	18
III.1.2.3. Les évolutions de SCOLA	18
III.2. Préparation des données.....	19
III.2.1. Création des jeux de données pour les analyses	19
III.2.1.1. Extraction et critères de choix pour les différentes analyses	19
III.2.1.2. Variables à l'étude	19
III.2.1.3. Agglomération des variables d'Activité économique.....	22
III.2.2. Traitement des non détectés	25
III.3. Modélisation statistique pour chaque agent chimique	25
III.3.1. Généralités sur les modèles linéaires mixtes	26
III.3.2. Modèle Tobit	27
III.3.2.1. Limites des modèles linéaires pour les données censurées	27

III.3.2.2.	Principe des modèles TOBIT	28
III.3.3.	Stratégie de modélisation	28
III.3.3.1.	La structure « Aléatoire »	30
III.3.3.2.	Modélisation de la structure « Nature de l'activité économique »	30
III.3.3.3.	Modélisation de la structure « Tendances temporelles et durée de prélèvement »	31
III.3.4.	Mesure de la qualité du modèle	33
III.3.5.	Interprétation des coefficients issus de la modélisation	33
III.4.	Méta analyse	34
III.4.1.	Principe général	34
III.4.2.	Application à notre étude	35
III.5.	Logiciels	36
CHAPITRE IV. Analyse descriptive et comparaison de deux bases de données d'exposition professionnelle françaises COLCHIC et SCOLA		37
IV.1.	Introduction	39
IV.2.	Methods	41
IV.2.1.	Origin of COLCHIC and SCOLA	41
IV.2.2.	Coding of the information in COLCHIC and SCOLA	41
IV.2.3.	Circumstances leading to records in COLCHIC and SCOLA	42
IV.2.4.	Descriptive summary of COLCHIC and SCOLA	45
IV.2.5.	Comparison of exposure levels between COLCHIC and SCOLA	46
IV.3.	Results	48
IV.3.1.	Descriptive summary of COLCHIC and SCOLA	48
IV.3.2.	Comparison of exposure levels between COLCHIC and SCOLA	55
IV.4.	Discussion	58
IV.4.1.	Descriptive comparison of COLCHIC and SCOLA	58
IV.4.2.	Comparison of exposure levels between COLCHIC and SCOLA	60
IV.4.3.	Limits of the study	61
IV.5.	Conclusion	62
CHAPITRE V. Analyse par modélisation de la base de données COLCHIC		63
V.1.	Introduction	63
V.2.	Méthode	64
V.2.1.	Préparation des données	64
V.2.2.	Variables exploitées	64
V.2.3.	Modélisation statistique pour chaque agent chimique	64

V.2.4. Méta analyse	65
V.3. Résultats	67
V.3.1. Analyse descriptive	67
V.3.2. Modélisation de COLCHIC	71
V.4. Discussion	78
V.4.1. Modélisation statistique	78
V.4.2. Limites de l'analyse.....	82
V.5. Conclusion.....	83
CHAPITRE VI. Modélisation comparative des bases de données COLCHIC et SCOLA .	84
VI.1. Introduction.....	84
VI.2. Méthode	85
VI.2.1. Préparation des données	85
VI.2.2. Variables exploitées.....	85
VI.2.3. Modélisation statistique pour chaque agent	86
VI.2.3.1. Analyse séparée de COLCHIC et SCOLA.....	86
VI.2.3.2. Analyse comparative de COLCHIC et SCOLA	86
VI.2.4. Méta analyse	87
VI.3. Résultats	89
VI.3.1. Analyse descriptive	89
VI.3.2. Structures des modèles.....	93
VI.3.3. Modélisation des bases de données d'Exposition professionnelle	95
VI.3.3.1. Modélisation séparée de COLCHIC et SCOLA.....	95
VI.3.3.2. Modélisation de la comparaison de COLCHIC et SCOLA.....	98
VI.3.3.3. Représentations graphiques des résultats des modèles.....	103
VI.4. Discussion	105
VI.4.1. Modélisation statistique des données d'exposition.....	105
VI.4.1.1. Analyse séparée de COLCHIC et SCOLA.....	105
VI.4.1.2. Analyse comparative de COLCHIC et SCOLA	107
VI.4.2. Limites de l'analyse.....	108
VI.5. conclusion.....	109
CHAPITRE VII. Discussion générale	110
VII.1. Modélisation statistique des données d'exposition	112
VII.1.1. Généralités	112
VII.1.2. COLCHIC 1987-2015, 2002-2015 et 2007-2015	112

VII.1.3.	SCOLA 2007-2015.....	115
VII.1.4.	Comparaison des données conjointes de COLCHIC et SCOLA 2007-2015.....	117
VII.2.	Utilisation des données de COLCHIC et SCOLA pour évaluer les niveaux de la population générale.....	118
VII.3.	Limites des analyses.....	120
VII.4.	originalité de ces travaux de recherche.....	121
VII.5.	Perspectives et recommandations.....	122
CHAPITRE VIII.	Conclusion générale	124

LISTE DES TABLEAUX

CHAPITRE I

Tableau I.1 Principales bases de données d'exposition : caractéristiques descriptives.....	6
Tableau I.2 Recommandations sur les informations devant accompagner les mesures dans les BDEP.....	7

CHAPITRE III

Tableau III.1 Principales variables documentées dans COLCHIC et SCOLA.....	23
--	----

CHAPITRE IV

Table IV.1 Main variables documented in COLCHIC and SCOLA (<i>Principales variables descriptives documentées dans Colchic et Scola</i>).....	43
Table IV.2 Distribution of records according to categories of the main variables in COLCHIC and SCOLA (<i>Distribution du nombre de mesures codées par catégorie de variables communes à COLCHIC et SCOLA</i>).....	50
Table IV.3 Descriptive statistics for personal measurements by chemical agent and sample duration in COLCHIC (1987-2012) and SCOLA (2007-2012) (<i>Statistiques descriptives des mesures d'exposition par agent chimique et par durée de prélèvement pour COLCHIC (1987-2012) et pour SCOLA (2007-2012)</i>)	53
Table IV.4 Comparison of exposure levels in classes of activity common to COLCHIC and SCOLA over the period 2007 to 2012 (<i>Comparaisons des niveaux d'exposition par classe d'activité commune à COLCHIC et SCOLA sur la période 2007 à 2012</i>)	56
Table IV.5 Comparison of exposure levels among selected chemical agents in establishments common to COLCHIC and SCOLA over the period 2007 to 2012 (<i>Comparaison des niveaux d'exposition des établissements communs à COLCHIC et SCOLA pour la période 2007 à 2012</i>).....	57

CHAPITRE V

Tableau V.1 Distribution du nombre de données (%) par catégories de variables incluses dans les modèles pour les périodes 1987-2015 et 2002-2015 par mesure individuelle et par mesure d'ambiance de travail.....	68
Tableau V.2 Choix de la structure « Activité économique ».....	72
Tableau V.3 Choix de la structure « Tendances temporelles et durée de prélèvement »	72
Tableau V.4 Coefficient de détermination par période et par type de prélèvement	73
Tableau V.5 Synthèse des indices d'exposition relative (RIE) et de leur intervalle de confiance à 95 % (IC 95 %) par Méta stratifiée par période 1987-2015 et 2002 – 2015 et par	

type de prélèvement.....	75
--------------------------	----

CHAPITRE VI

Tableau VI.1 Distribution du nombre de données individuelles (%) par catégorie de variables incluses dans les modèles pour la période 2007-2015 pour COLCHIC, SCOLA et la BDEP commune (Comp. C&S).....	90
---	----

Tableau V.2 Choix de la structure « Activité économique ».....	94
--	----

Tableau V.3 Choix de la structure « Tendances temporelles et durée de prélèvement ».....	94
--	----

Tableau V.4 Coefficient de détermination par période et par type de prélèvement.....	95
--	----

Tableau VI.5 Synthèse des indices d'exposition relative (RIE) et de leur intervalle de confiance à 95 % (IC 95 %) par Méta analyse pour COLCHIC et SCOLA sur la période 2007-2015.....	97
--	----

Tableau VI.6 Synthèse des indices d'exposition relative (RIE) et de leur intervalle de confiance à 95 % (IC 95 %) par Méta analyse pour la BDEP commune avec les interactions entre la « Source » et les autres variables sur la période 2007-2015.....	99
---	----

Tableau VI.7 Synthèse des indices d'exposition relative (RIE) par Méta analyse pour COLCHIC, SCOLA et la BDEP commune avec les interactions entre la « Source » et les autres variables sur la période 2007-2015.....	101
---	-----

Tableau VI.8 Ratios SCOLA/COLCHIC pour les 4 scénarios et de leur intervalle de confiance à 95 % (IC 95 %) par Méta analyse	102
---	-----

CHAPITRE VII

Tableau VII.1 : Synthèse des principaux résultats obtenus lors des analyses des bases de données COLCHIC et SCOLA.....	111
--	-----

LISTE DES FIGURES

CHAPITRE I

Figure I.1 : Schéma conceptuel des biais dans une base de données d'exposition professionnelle.....	9
---	---

CHAPITRE IV

Figure IV.1 Distribution of the number of data per year (Distribution du nombre de mesures par année).....	48
--	----

Figure IV.2 Distribution of n for the Industrial sector and comparison with the population by industrial sector in 2012 (Distribution du nombre de mesures par secteur d'activité et comparaison avec le nombre de travailleurs (en 2012) de ces mêmes secteurs industriels).....	51
---	----

CHAPITRE V

Figure V.1 Distribution annuelle du nombre de mesures individuelles et d'ambiance dans COLCHIC.....	67
---	----

Figure V.2 : RIE issus de la Méta analyse par agent chimique mesuré en individuel sur la période 2002-2015 pour les catégories : A- Equipement de protection individuelle adapté (Ref : non adapté) B- Type de procédé fermé (Ref : ouvert) C- Exposition fréquente (Ref : permanent).....	77
--	----

CHAPITRE VI

Figure VI.1 Distribution annuelle du nombre de mesures individuelles dans COLCHIC, SCOLA et la BDEP commune sur la période 2007-2015.....	89
---	----

Figure VI.2 : Ratios SCOLA/COLCHIC issus de la Méta analyse par agent chimique pour les 4 scénarios : A- 2007 et durée du prélèvement de 30 minutes, B- 2015 et durée du prélèvement de 30 minutes, C- 2007 et durée du prélèvement de 240 minutes et D- 2015 et durée du prélèvement de 240 minutes.....	104
---	-----

LISTE DES ANNEXES

CHAPITRE IV

ANNEXE IV.1 : MEASUREMENT DISTRIBUTION CURVES BY DATABASE FOR THE PERIOD 2007 – 2012 (COURBES DE DISTRIBUTION CUMULEE DES EXPOSITIONS PAR AGENT CHIMIQUE POUR LA PERIODE 2007 A 2012 POUR COLCHIC ET SCOLA).....	131
ANNEXE IV.2 : COMPARISON OF EXPOSURE LEVELS IN CLASSES OF ACTIVITY COMMON TO COLCHIC AND SCOLA OVER THE PERIOD 2007 TO 2012 RESTRICTED TO THE SAMPLE DURATION BETWEEN 240 AND 480 MINUTES (COMPARAISON DES NIVEAUX D'EXPOSITION PAR CLASSE D'ACTIVITE COMMUNE A COLCHIC ET SCOLA SUR LA PERIODE 2007 A 2012 ET RESTREINTE AUX DUREES DE PRELEVEMENT COMPRISES ENTRE 240 ET 480 MINUTES).....	134

CHAPITRE V

ANNEXE V.1 POURCENTAGE DE DONNEES INFERIEURES A LA LIMITE DE QUANTIFICATION, REPARTITION DE LA PROPORTION DE MESURES COURTE DUREE (CD) ET LONGUE DUREE (LD) PAR AGENT CHIMIQUE COMPTANT PLUS DE 1 000 RESULTATS SUR LA PERIODE 1987-2015 POUR LES MESURES INDIVIDUELLES ET LES MESURES D'AMBIANCE DE TRAVAIL.....	135
ANNEXE V.2 : POURCENTAGE DE DONNEES INFERIEURES A LA LIMITE DE QUANTIFICATION, REPARTITION DE LA PROPORTION DE MESURES COURTE DUREE (CD) ET LONGUE DUREE (LD) PAR AGENT CHIMIQUE COMPTANT PLUS DE 1 000 RESULTATS SUR LA PERIODE 2002-2015 POUR LES MESURES INDIVIDUELLES ET LES MESURES D'AMBIANCE DE TRAVAIL.....	138
ANNEXE V.3 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) ET DE LEUR INTERVALLE DE CONFIANCE A 95 % (IC 95 %) POUR LES PRELEVEMENTS INDIVIDUELS PAR META ANALYSE STRATIFIEE PAR PERIODE 1987-2015 ET 2002–2015 ET PAR FAMILLE D'AGENTS CHIMIQUES.....	140
ANNEXE V.4 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) ET DE LEUR INTERVALLE DE CONFIANCE A 95 % (IC 95 %) POUR LES PRELEVEMENTS D'AMBIANCE PAR META ANALYSE STRATIFIEE PAR PERIODE 1987 - 2015 ET 2002 – 2015 ET PAR FAMILLE D'AGENTS CHIMIQUES	142
ANNEXE V.5 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) ET DE LEUR INTERVALLE DE CONFIANCE A 95 % (IC 95 %) POUR LES PRELEVEMENTS INDIVIDUELS PAR META ANALYSE DES MODELES AVEC ET SANS VARIABLE ALEATOIRE STRATIFIEE PAR PERIODE 1987-2015 ET 2002 – 2015	144
ANNEXE V.6 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) ET DE LEUR INTERVALLE DE CONFIANCE A 95 % (IC 95 %) POUR LES PRELEVEMENTS D'AMBIANCE PAR META ANALYSE DES MODELES AVEC ET SANS VARIABLE ALEATOIRE STRATIFIEE PAR PERIODE 1987-2015 ET 2002–2015.....	146
ANNEXE V.7 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR META ANALYSE DES AGENTS MESURES EN INDIVIDUEL SUR UNE LONGUE DUREE (LD) ASSOCIEE AUX PRELEVEMENTS COURTE DUREE SI LEUR PROPORTION EST SUPERIEURE A 20 % (CD>20 %) ET PAR PERIODE 1987-2015 ET 2002–2015.....	148
ANNEXE V.8 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR META ANALYSE DES AGENTS MESURES EN AMBIANCE SUR UNE LONGUE DUREE (LD) ASSOCIEE AUX PRELEVEMENTS COURTE DUREE SI LEUR PROPORTION EST SUPERIEURE A 20 % (CD>20 %) ET PAR PERIODE 1987-2015 ET 2002–2015.....	150

CHAPITRE VI

ANNEXE VI.1 POURCENTAGE DE DONNEES INDIVIDUELLES INFERIEURES A LA LIMITE DE QUANTIFICATION, REPARTITION DE LA PROPORTION DE MESURES COURTE DUREE (CD) ET LONGUE DUREE (LD) PAR AGENT CHIMIQUE COMPTANT PLUS DE 1 000 RESULTATS SUR LA PERIODE 2007-2015 POUR COLCHIC, SCOLA ET LA COMPARAISON COLCHIC SCOLA	152
ANNEXE VI.2 COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR META ANALYSE DE COLCHIC POUR LES FAMILLES D'AGENTS CHIMIQUES ET DES MODELES AVEC ET SANS VARIABLE ALEATOIRE	155
ANNEXE VI.3 COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR META ANALYSE DE SCOLA POUR LES SOLVANTS ET DES MODELES AVEC ET SANS VARIABLE ALEATOIRE	157
ANNEXE VI.4 COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR META ANALYSE POUR LA COMPARAISON DE COLCHIC ET SCOLA SANS INTERACTION ET AVEC INTERACTIONS POUR TOUS LES SECTEURS D'ACTIVITE ET RESTREINT AU SECTEURS COMMUNS AUX BDEP	159
ANNEXE VI.5 EVOLUTION DES RATIOS SCOLA/COLCHIC.....	161

ABRÉVIATIONS

ACGIH American Conference of Governmental Industrial Hygienists

AIHA American Industrial Hygiene Association

AIC Akaike Information Criterion

ANOVA Analysis of variance (*Analyse de la Variance*)

ANSES Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail

BDEP Base de Données d'Exposition Professionnelle

BIA German Institute for Occupational Safety

IFA BG Institut D'hygiène et de Sécurité / Institut pour la Sécurité et la Santé de l'assurance Sociale Allemande des Accidents, Allemagne

BIC Bayesian Information Criterion

CAPPROEX Catégorie de Produit ou de Procédé à l'Origine de l'Exposition Professionnelle

CARSAT Caisse d'Assurance Retraite et de la Santé au Travail

CAS Chemical Abstracts Service

CAT-MP Commission Accidents du Travail et Maladies Professionnelles

CITP Classification Internationale Type de Professions

CMR Cancérogène, Mutagène ou Reprotoxique

CNAMTS Caisse Nationale d'assurance Maladie des Travailleurs Salariés

COLCHIC Système de Collecte des Données d'exposition Chimiques des Laboratoires des Caisses Régionales d'assurance Maladie

COLPHY Système de Collecte de données techniques sur les risques physiques et la ventilation

COV Composé Organique Volatile

CRAMIF Caisse Régionale d'Assurance Maladie d'Ile De France

DE Détecté

DGT Direction Générale du Travail

ECDF Empirical Cumulative Distribution Curves

GM Geometric Mean

HSE Health And Safety Executive

IMIS Integrated Management Information System

INGRES Interactions Gènes – Risques Environnementaux et Effets sur la Santé

INRS Institut National de Recherche et de Sécurité

INSEE Institut National de la Statistique et des Etudes Economiques

INSERM Institut National de la Santé et de la Recherche Médicale

ISPESL Institut Italien pour la Sécurité et la Prévention

LIC Laboratoires Interrégionaux de Chimie

LIMS Laboratory Information Management System

LoD Limit of detection (*Limite de détection*)

LoQ Limit of quantification (*Limite de quantification*)

META Microscopie Électronique à Transmission Analytique

MOCP Microscopie Optique à Contraste de Phase

NACE Nomenclature statistique des Activités économiques dans la Communauté Européenne

NAF Nomenclature des Activités française

OELV Occupational Exposure Limit Value

OSHA Occupational Safety and Health Administration

REML Restricted Maximum Likelihood

RIE Relative indice of exposure (*Indice d'Exposition Relative*)

ROME Répertoire Opérationnel des Métiers et des Emplois

SCOLA Système de COLlecte des Laboratoires Accrédités

TEXAS Tools EXposure ASsessment

VLEP Valeur Limite d'Exposition Professionnelle

VLEP-CT Valeur Limite d'Exposition Professionnelle de courte durée

ORGANISATION DE LA THÈSE

Cette thèse, divisée en 8 chapitres, présente aux chapitres I et II la problématique générale de l'utilisation des bases de données d'exposition professionnelle (BDEP) et des bases de données françaises COLCHIC et SCOLA pour l'évaluation de l'exposition professionnelle puis les objectifs de travaux de thèse. Le chapitre III décrit les méthodes générales utilisées, dont la sélection des données et des prédicteurs dans COLCHIC et SCOLA, la technique de modélisation statistique employée ainsi que la méthode de Méta analyse utilisée pour synthétiser les résultats à travers les agents. Le cœur de la thèse est formé par les chapitres IV à VI qui portent tout d'abord, sur l'analyse descriptive et comparative de COLCHIC et SCOLA publiée dans American Journal of Industrial Medicine puis, sur la modélisation de COLCHIC et enfin, sur la modélisation comparative des bases de données COLCHIC et SCOLA. Les chapitres VII et VIII présentent la discussion générale des résultats et les conclusions tirées de ce travail.

CHAPITRE I. INTRODUCTION GÉNÉRALE

I.1. CONTEXTE

La connaissance *a priori* de l'environnement des postes de travail et des conditions d'exposition des opérateurs aux agents chimiques joue un rôle essentiel en santé au travail pour plusieurs raisons. Lors de l'évaluation du risque chimique, le recensement, de manière exhaustive, des agents chimiques présents sur le lieu de travail est un enjeu important. L'emploi de données historiques existantes permet de faciliter cet exercice [Begin et al., 1995]. Cette connaissance permet la mise en place de politiques de prévention, ciblées par secteur d'activité ou par métier. De plus, elle présente un intérêt important en tant qu'outil pour la conduite d'études épidémiologiques (en particulier pour les agents ayant des effets à long terme comme les cancérogènes), pour l'analyse de risques toxicologiques ou encore pour la mise en place de modèles d'aide à la décision pour l'évaluation des expositions professionnelles, par exemple, l'outil TEXAS (Tools EXposure ASsessment), actuellement élaboré à l'Institut National de Recherche et Sécurité, INRS [Clerc et al., 2015].

I.2. LES BASES DE DONNÉES D'EXPOSITION PROFESSIONNELLE

Dans le but de centraliser les informations d'expositions professionnelles suite à la collecte d'échantillons d'air, de produits..., plusieurs pays ont construit des bases de données informatisées. Parmi celles-ci, nous pouvons citer, comme exemple, la base SIREP sur les expositions à des agents cancérogènes en Italie [Scarselli et al., 2007 ; Scarselli, 2011], la base NEBD au Royaume-Uni [Burns et al., 1989], la base IMIS aux Etats-Unis [Henn et al., 2011 ; Lavoué et al., 2012], la base MEGA en Allemagne [Stamm, 2001 ; Koppisch et al., 2012], le registre des mesures d'hygiène du travail en Finlande [Kauppinen, 2001] ou encore les bases COLCHIC et SCOLA en France [Carton et Goberville, 1989 ; Vincent et Jeandel, 2001 ; Mater et al., 2016]. Ces bases peuvent être utilisées dans différents buts. Dans certains cas, elles sont employées à des fins de traçabilité des informations recueillies sur le terrain par les préventeurs, inspecteurs, médecins... et par le laboratoire qui analyse les échantillons prélevés ou encore pour

l'archivage des expositions mesurées en milieu industriel. COLCHIC est, par exemple, utilisée par les huit laboratoires interrégionaux de chimie (LIC) des Caisses d'Assurance Retraite et de la Santé Au Travail (CARSAT) et de la Caisse Régionale d'Assurance Maladie d'Ile de France (CRAMIF). Ainsi, les mesures d'exposition collectées en entreprise sont enregistrées puis archivées dans les bases de données d'exposition professionnelle (BDEP). Dans le cadre de leur exploitation, elles peuvent être utilisées pour aider à la surveillance des expositions professionnelles et ainsi cibler les situations à traiter en priorité. Dans d'autres cas, elles sont employées à des fins de soutien à la réalisation d'études épidémiologiques (par exemple, définir le niveau de risque d'une population exposée à un agent cancérigène, mutagène ou reprotoxique (CMR)) rétrospectives (par exemple, pour aider au constat d'une déclaration de maladie professionnelle), prospectives (par exemple, comme un outil d'aide à l'évaluation du risque chimique) et/ou toxicologiques (par exemple, pour contribuer à l'élaboration d'une valeur limite d'exposition professionnelle). Le Tableau 1 présente une liste des principales bases de données nationales existantes, ainsi que leur année de création, le propriétaire et leurs objectifs.

Chaque base a une structure qui lui est propre. Le recueil et la codification des informations qui accompagnent chaque mesure doivent répondre, dans un souci d'homogénéité et d'harmonisation des éléments intégrés au sein de la base, à des standards définis au préalable. Ainsi, les tables de référence utilisées peuvent varier d'une base à une autre : certaines sont de références internationales, d'autres nationales ou propres à la base elle-même. Plusieurs auteurs ou organisations ont émis des recommandations concernant le degré d'information devant accompagner les mesures d'exposition. Deux efforts notables ont été présentés par des groupes de travail européen et américain formés au milieu des années 1990 [ACGIH-AIHA, 1996 ; Rajan et al., 1997]. Le Tableau I.2 présente les grandes catégories d'informations jugées importantes par le groupe européen ainsi qu'une brève description de leur contenu. Les recommandations américaines étaient sensiblement similaires.

Tableau I.1 Principales bases de données d'exposition : caractéristiques descriptives

Pays	Nom	Année de création	Propriétaire / Gestionnaire	Qui renseigne la base ?	Nombre de mesures	Nombre de substances mesurées	Principaux objectifs
Italie	SIREP	1996	ISPESL ¹	Les institutions scientifiques et gouvernementales	> 250 000	550	• Prévention • Evaluation, contrôle et réduction du risque cancérogène
Royaume-Uni	NEDB	1986	HSE ²	Inspecteurs du HSE	Quelques dizaines de milliers	400	• Prévention • Elaboration de politiques réglementaires
Etats-Unis	IMIS	1971	OSHA ³	Inspecteurs de l'OSHA	> 1 900 000	1 082 ⁴	• Prévention • Vérification de la conformité avec la réglementation
Allemagne	MEGA	1972	BIA / IFA ⁵	Inspecteurs du BGs (Instituts d'assurances et de prévention)	2 200 000	809 ⁶	• Prévention • Assurance
France	COLCHIC	1987	CNAMTS / INRS	LIC des CARSAT/CRAM et agents de l'INRS	950 000	700	• Prévention
France	SCOLA	2007	DGT / INRS	Organismes de contrôle accrédités selon NF EN ISO 17025	435 000	105	• Prévention • Réglementation

¹ Institut supérieur pour la prévention et la sécurité du travail, Italie

² Health and Safety Executive, Royaume-Uni

³ Occupational Safety and Health Administration, États-Unis

⁴ [Lavoué et al., 2012]

⁵ BG Institut d'hygiène et de sécurité / Institut pour la sécurité et la santé de l'assurance sociale des accidents, Allemagne

⁶ [Koppisch et al., 2012]

Tableau 1.2 Recommandations sur les informations devant accompagner les mesures dans les BDEP

Appellation de la catégorie	Types d'informations incluses
Situation	Description de l'entreprise (activité économique, taille, adresse)
Aire de travail	Description du lieu de mesure (département, poste de travail, procédé utilisé)
Activité de l'employé	Description de la profession et des tâches réalisées par l'employé
Produit	Identification du produit / matière première / intermédiaire de production qui est la source d'exposition
Agent chimique	Description de l'agent chimique mesuré (numéro CAS, noms commerciaux)
Éléments influents sur l'exposition	Informations sur des paramètres déterminants comme la structure de l'exposition (continue, intermittente), les moyens de maîtrise (ventilation locale / générale) ou encore les équipements de protection individuelle
Stratégie de mesure	Informations sur les objectifs de la mesure (évaluation du pire des cas pour la conformité aux normes, établissement d'un portrait représentatif, évaluation des moyens de maîtrise de l'exposition)
Procédure de mesurage	Informations sur la méthode d'échantillonnage et d'analyse (méthode analytique, durée d'échantillonnage par rapport à la durée d'exposition, date d'échantillonnage)
Résultats	Valeur numérique de la mesure
Référence	Référence du rapport d'hygiène correspondant à la mesure

Les acteurs alimentant les BDEP proviennent d'activités différentes : inspecteurs (IMIS, MEGA, NEDB) chargés de vérifier la conformité aux normes du travail, préventeurs (COLCHIC), personnels d'organismes de contrôles privés (SCOLA)...

L'accès aux données enregistrées dans les BDEP varie en fonction des clauses de confidentialité régies par le propriétaire de la base. Par exemple, le système intégré de gestion de l'information (IMIS), créé par l'Occupational Safety and Health Administration (US-OSHA) en 1972, donne libre accès au grand public aux informations enregistrées pour chaque entreprise par les inspecteurs fédéraux de l'OSHA ou de l'état. A contrario, les informations collectées par les utilisateurs de COLCHIC et SCOLA sont uniquement accessibles après une exploitation statistique dans des rapports de recherche publics.

I.2.1. LES LIMITES DE L'UTILISATION DES BASES DE DONNÉES NATIONALES COMME SOURCES D'INFORMATION SUR L'EXPOSITION

Malgré le potentiel important d'informations sur l'exposition professionnelle, représenté par les bases de données décrites précédemment, il est nécessaire de garder à l'esprit qu'aucune d'entre-elles n'a été conçue avec comme objectif d'être représentative de la population des travailleurs en général. Ainsi, chaque étape conduisant à la mesure puis à l'enregistrement d'un niveau d'exposition dans une BDEP est susceptible d'introduire des biais dans l'interprétation de cette mesure par rapport à la population générale [Olsen et al., 1991].

I.2.2. ORIGINE DES BIAIS POTENTIELS

Les biais potentiellement introduits dans les BDEP peuvent être conceptualisés de manière simple et sont illustrés dans la Figure 1.1 [Lavoué, (2006), Évaluation de l'exposition professionnelle au formaldéhyde à partir de sources de données préexistantes. Thèse de doctorat], reproduite avec l'autorisation de l'auteur.

Parmi les biais présentés Figure 1.1, certains peuvent être éliminés si les informations caractérisant les mesures sont suffisamment renseignées dans la BDEP. Prenons l'exemple de l'industrie de l'héliogravure, décrit par Clerc et Vincent [2014], où il est montré que le métier de receveur présente des niveaux de concentration au toluène bien inférieurs au métier de conducteur (moyenne géométrique respectivement de 44,18 mg.m⁻³ et 101,72 mg.m⁻³). Dans ce contexte, si les mesures avaient été enregistrées sans information du métier dans la BDEP, un biais aurait alors été généré sur le niveau général d'exposition des travailleurs du secteur de l'héliogravure. Dans le cadre de cette étude, l'enregistrement du métier dans COLCHIC a permis d'éliminer ce biais.

D'autres biais, liés par exemple, à la stratégie d'échantillonnage, sont cependant plus compliqués à appréhender. Prenons un exemple, toujours dans le secteur de l'héliogravure, Clerc et Vincent [2014] ont montré dans leur étude que les niveaux de concentration mesurés sont liés à la durée de prélèvement. Plus la durée est courte, plus les niveaux mesurés sont importants. Bien que réaliser une mesure de plus courte durée permette dans certaines situations la maîtrise des conditions d'exposition, elle ne sera

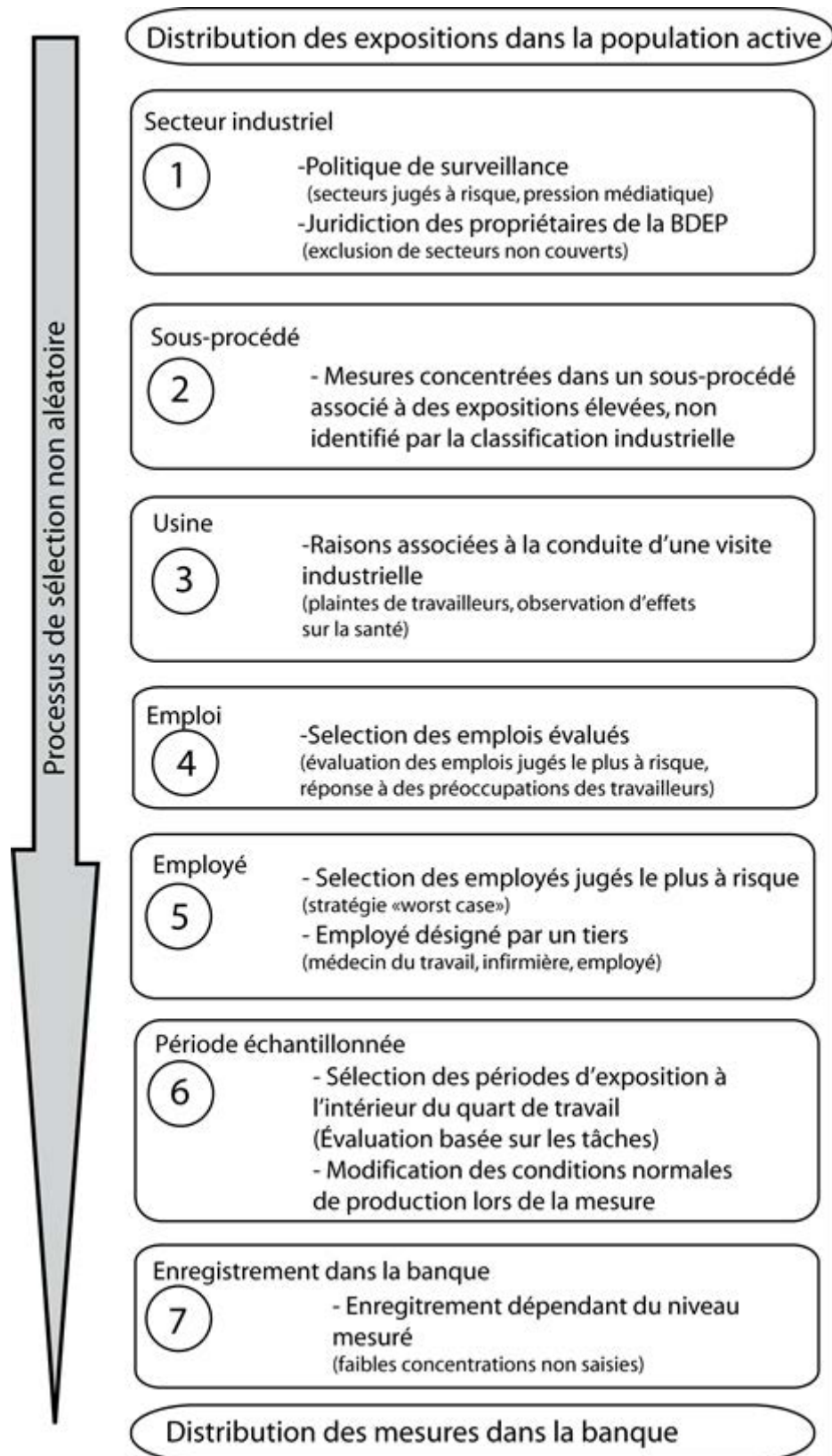


Figure I.1 : Schéma conceptuel des biais dans une base de données d'exposition professionnelle [Lavoué, (2006), Évaluation de l'exposition professionnelle au formaldéhyde à partir de sources de données préexistantes. Thèse de doctorat]

pas, globalement, représentative de l'exposition moyenne du travailleur. Des études spécifiques pour analyser ce type de biais sont alors nécessaires [Olsen, 1996]. A titre d'exemple, Lavoué et al. [2005] ont identifié, dans les usines de fabrication de panneaux de bois reconstitués, une différence systématique entre les niveaux d'exposition au formaldéhyde mesurés au Québec par une équipe de recherche et par des hygiénistes d'agences gouvernementales. Pour leurs prédictions, les chercheurs ont corrigé leurs estimations afin de tenir compte de cette observation. Il en a été de même lors de l'analyse du formaldéhyde dans COLCHIC et IMIS. Lavoué et al. [2010], après avoir observé cette association, ont aussi pu corriger leurs prédictions.

L'étude des différences entre l'exposition réelle dans la population et les niveaux enregistrés dans une BDEP ne peut, en général, être réalisée directement, puisqu'on ne dispose presque jamais de résultats de campagnes de mesures aléatoires de grande ampleur. L'étude des biais peut être abordée selon deux angles : l'association entre les niveaux enregistrés et certaines variables internes ou la comparaison avec des sources de données externes. Ainsi, les BDEP contiennent souvent plusieurs variables descriptives du milieu de travail visité telles que la taille de l'établissement, le secteur d'activité... Si la distribution des tailles d'entreprises pour un secteur d'activité diffère dans la BDEP de celle de la population générale, un biais sera alors généré. De même, lorsque des variables mettent en avant des approches différentes dans le choix des situations à évaluer (par exemple, visite d'un préventeur / visite associée à un cadre réglementaire), le constat d'une variation des niveaux enregistrés dans ces deux situations sera indicateur d'un biais et permettra de l'évaluer. Dans ce cas, l'expert devra choisir quelle situation lui semble être la plus proche de la population générale. Pour conclure, les écarts entre les différentes stratégies employées peuvent être mesurés en comparant la BDEP avec d'autres sources d'information.

I.2.3. BIAIS RAPPORTÉS DANS LA LITTÉRATURE

Globalement, il apparaît que peu d'études se sont attachées à étudier de façon approfondie les biais présents dans les bases de données d'exposition professionnelle. Une revue récente montre que la base états-unienne IMIS a été la plus fréquemment étudiée dans la littérature scientifique, avec plus de 20 articles publiés, dont certains

abordent la question de la représentativité [Lavoué et al., 2012]. Les auteurs ont conclu de ce travail que les résultats disponibles, principalement obtenus à l'aide d'analyses internes, n'ont pas montré la présence de biais notables de façon répétée. Cependant, des exercices de validation externe ont montré des biais liés, par exemple, à la stratégie de prélèvement [Olsen et al., 1991]. Dans cette étude, il a été montré que pour une même population, les différentes stratégies déployées engendraient des variations d'un facteur 5 à 10 sur les concentrations mesurées en toluène. De même, lors de la comparaison des niveaux d'exposition de xylène enregistrés dans 5 bases de données européennes, Vinzents et al. [1995] ont conclu à l'existence d'un biais modéré lié à la stratégie mise en place (choix des entreprises, des travailleurs et durée d'échantillonnage). Enfin, deux autres exemples, l'un entre IMIS et une base similaire intégrant des résultats associés à une stratégie préventive plutôt que de contrôle (agent chimique : plomb) [Okun et al., 2004], et l'autre comparant des bases européennes et une base canadienne (agent chimique : silice cristalline) [Peters et al., 2011], ont aussi mis en évidence des différences systématiques, bien que de faible amplitude.

En ce qui concerne les bases françaises, lors de l'analyse des niveaux d'exposition au formaldéhyde dans COLCHIC, Lavoué et al. [2005] ont observé des concentrations sur 8-heures près de 2 fois supérieures (résultats non statistiquement significatifs) lors des évaluations faisant suite à l'identification d'exposition possible, par comparaison aux enquêtes systématiques. Ces résultats suggèrent une différence possible de stratégie de mesure entre les deux types d'activité. Les auteurs concluaient que ces observations nécessitaient une évaluation systématique pour d'autres agents.

En conclusion, les BDEP représentent un fort potentiel d'information sur l'exposition professionnelle, qui peut être utilisé pour améliorer plusieurs aspects de la prévention des maladies professionnelles. La France est très bien positionnée dans ce domaine puisque deux bases de taille importante y sont présentes, COLCHIC et SCOLA. Cependant, la représentativité des mesures des conditions de travail réelles et historiques en France qu'elles contiennent est largement inconnue. Il apparaît donc nécessaire d'identifier et de caractériser les biais potentiels présents dans ces deux bases pour améliorer l'interprétation des résultats qui y sont stockés et leur utilisation à des fins de prévention.

CHAPITRE II. OBJECTIFS DES TRAVAUX DE THÈSE

II.1. OBJECTIF GÉNÉRAL

La finalité de ces travaux consiste en l'utilisation des données de COLCHIC et SCOLA pour la prévention des maladies professionnelles, incluant la surveillance des expositions, l'élaboration d'outils d'analyse de risque, et l'estimation rétrospective de l'exposition pour l'épidémiologie. Cet objectif vise à répondre à la question : « *Dans quelle mesure les données d'exposition à des produits chimiques, enregistrées dans les bases COLCHIC et SCOLA, reflètent les expositions professionnelles de façon historique en France ?* » et ainsi permettre d'interpréter au mieux les données de ces bases pour leurs différentes utilisations.

II.2. OBJECTIFS SPÉCIFIQUES

L'opérationnalisation de cette étude comprend les objectifs spécifiques suivants :

- Identifier dans COLCHIC et SCOLA des variables associées de façon systématique aux niveaux mesurés.

D'une part, certaines variables sont associées aux stratégies employées pour effectuer les mesures, comme par exemple, le motif de la demande documentée dans COLCHIC alors qu'il est systématiquement le même dans SCOLA. La mise en évidence d'une relation entre cette variable et les niveaux d'exposition mesurés permettrait, après avoir défini le motif le plus susceptible de donner lieu à des mesures représentatives de l'exposition, de corriger les estimations finales. D'autre part, des variables peuvent se révéler associées aux niveaux d'exposition et se distribuer différemment dans COLCHIC/SCOLA et dans la population générale (par exemple, la taille des établissements). Là encore, cette approche permettrait de corriger des estimations issues des BDEP. Finalement pour les prédicteurs communs, des différences dans les associations entre les deux bases de données suggèreraient un effet global de stratégie d'échantillonnage.

- Identifier des différences systématiques entre les bases COLCHIC et SCOLA en tenant compte des associations préalablement mises en évidence.

La présence en France de deux BDEP contenant des résultats provenant du même tissu industriel, mais mesurés avec des objectifs différents, représente un terrain privilégié. Après la prise en compte des variables identifiées comme influentes lors de l'analyse initiale de chaque base, l'observation ou non de différences systématiques entre les niveaux dans les deux bases, pour des activités industrielles similaires, pourrait nous informer sur leur représentativité. Des niveaux similaires dans les deux bases tendraient à soutenir leur utilisation directe et combinée comme source d'information sur l'exposition. Des niveaux systématiquement différents demanderaient de porter un jugement sur laquelle des deux sources serait la plus représentative et sur une pondération éventuelle des niveaux estimés à partir des deux bases.

Afin de répondre à ces objectifs spécifiques, 3 tâches ont été entreprises : tout d'abord, une analyse descriptive précise du contenu de COLCHIC et SCOLA pour définir le cadre des analyses approfondies ensuite, l'identification séparée dans COLCHIC et SCOLA des variables influentes sur les niveaux d'exposition pour une large gamme d'agents chimiques par le biais de modèles statistiques et enfin, la comparaison de ces deux BDEP après prise en compte des variables identifiées comme influentes dans les analyses séparées.

CHAPITRE III. MÉTHODES

Ce chapitre a pour objectif de présenter 1- les bases de données COLCHIC et SCOLA 2- la démarche générale entreprise pour préparer les données 3- le choix des modèles statistiques utilisés pour identifier les facteurs associés aux niveaux d'exposition pour chaque agent 4- la méthode de Méta analyse utilisée pour la synthèse des résultats de modélisation à travers les agents. Nous noterons que les spécificités propres à certaines analyses de modélisation ne sont pas abordées dans le présent chapitre, mais sont décrites dans le paragraphe « Méthode » des chapitres concernés.

III.1. PRÉSENTATION DES BASES DE DONNÉES D'EXPOSITION PROFESSIONNELLE FRANÇAISES

L'INRS administre plusieurs bases de données d'exposition professionnelle. Parmi elles : 1- COLPHY, déployée en 2016, enregistre des informations techniques sur les risques physiques et la ventilation 2- COLCHIC et SCOLA portent sur les expositions aux produits chimiques et/ou biologiques dans les entreprises françaises.

III.1.1. COLCHIC

III.1.1.1. Généralités

Créée en 1986 à la demande de la commission des accidents du travail et des maladies professionnelles (CAT-MP), la base de données française COLCHIC [Carton et Goberville, 1989 ; Carton et Jeandel, 1993 ; Vincent et Jeandel, 2001 ; Lavoué et al., 2011 ; Mater et al., 2016] est alimentée par des mesures sur des échantillons de produits industriels et par des mesures d'exposition aux agents chimiques et biologiques effectuées par les LIC des CARSAT / CRAM et par les laboratoires de l'INRS. Sa mise en place a permis de répondre à plusieurs objectifs :

- ▶ La centralisation des mesures d'exposition ;
- ▶ L'harmonisation des pratiques techniques de métrologie des atmosphères de lieux de travail en suivant les procédures définies par l'institution dans la base MétroPol ;

- L'exploitation des informations enregistrées afin d'identifier et de définir les axes prioritaires de prévention en France.

Sur la période 1987 à 2015, COLCHIC compte 114 agents chimiques pour chacun desquels plus de 1 000 données dans l'air des lieux de travail ont été renseignées. Les données de cette base sont utilisées principalement pour des opérations d'évaluation de risque, menées au niveau national et parfois au niveau international. Les informations collectées dans COLCHIC sont généralement codifiées à l'aide :

- D'une classification internationale : l'identification des substances dans COLCHIC est effectuée grâce à leur numéro d'enregistrement unique, le numéro CAS (Chemical Abstracts Service) ;
- D'une classification nationale : la codification du secteur d'activité de l'entreprise visitée ou du métier du travailleur sur lequel le prélèvement est effectué repose respectivement sur la Nomenclature des Activités Françaises (NAF révision 2, 2008) et sur le Répertoire Opérationnel des Métiers et des Emplois (ROME version 1.2, 1999). Gérée par l'Institut National de la Statistique et des Etudes Economiques (INSEE), la nomenclature statistique nationale d'activité est directement emboîtée dans la nomenclature d'activités européenne (NACE), mais comporte un niveau de classification en plus (4 chiffres et une lettre). Le code ROME, géré par Pôle Emploi, est structuré en 4 niveaux (les catégories professionnelles, les domaines professionnels, les emplois/métiers et les différentes appellations de métiers et d'emplois).
- D'un système propre à COLCHIC : la tâche effectuée par le travailleur durant sa journée de travail est codée dans COLCHIC à partir d'une table de référence qui lui est propre. Cette dernière a été élaborée par un groupe de travail composé de représentants des LIC et de l'INRS. D'autres informations telles que la méthode de prélèvement ou d'analyse, l'objectif de la mesure, sa représentativité, la ventilation globale, les équipements de protection collective... sont aussi codées dans COLCHIC à partir de tables de référence internes.

Pour les autres informations, les utilisateurs les renseignent dans un format alpha numérique : par exemple, pour saisir un numéro d'échantillon, l'utilisateur utilisera la

référence propre au système de codification du laboratoire.

III.1.1.2. Exploitation des données de COLCHIC

Des procédures de requêtes, accessibles uniquement aux LIC et à l'INRS, permettent d'extraire anonymement, les informations enregistrées dans COLCHIC en fonction de critères de recherche tels que le secteur d'activité, le métier, la tâche, la période d'exposition... Ces données peuvent alors être utilisées dans le cadre de travaux de recherche portant sur 1- des expositions à différents produits chimiques : plomb [Carton 1995], formaldéhyde [Carton, 1995 ; Vincent et Jeandel, 2002 ; Lavoué et al., 2006], fibres minérales [Kauffer et Vincent, 2007], toluène [Clerc et al., 2015a] 2- des modèles de prédiction des expositions [Clerc et al., 2015b].

COLCHIC est régulièrement sollicitée par des organismes tels que le Ministère du Travail, l'ANSES, l'agence Santé Public France (anciennement InVS), l'Inserm, les CARSAT... Son exploitation permet, en particulier, d'apporter de la connaissance sur les conditions d'exposition lors de la révision des tableaux de Maladies Professionnelles ou lors d'évaluations des risques menées par l'ANSES (par exemple, fibres céramiques réfractaires, formaldéhyde...). Au total, depuis sa mise en production, plus de 500 exploitations statistiques à façon ont été menées par l'INRS.

Pour faciliter la mise à disposition des informations, deux applications, SOLVEX et FIBREX, portant respectivement sur les composés organiques volatils et sur les fibres (hors amiante), ont été créées puis déployées sur le site de l'INRS. Elles permettent, à partir de critères renseignés par l'utilisateur pouvant porter sur le secteur d'activité, le métier ou la substance et pour une période définie, d'obtenir des données statistiques sur les niveaux d'exposition *a priori*. Les informations issues de SOLVEX et FIBREX proviennent des données enregistrées dans COLCHIC.

III.1.1.3. Les évolutions de COLCHIC

En 2002, COLCHIC 2 est déployée. Un nouveau système de codification est mis en place [Vincent et Jeandel, 2001 ; Mater et al., 2016] : création de 6 nouvelles variables (le métier, l'activité de production estimée, le type de procédé, la fréquence d'exposition, le code

CAPPROEX : catégorie de produit ou de procédé à l'origine de l'exposition professionnelle et la température d'utilisation du produit) et modification/évolution de 5 variables existantes (l'objectif, la représentativité, le type de lieu de travail, la ventilation globale et la protection collective).

En 2010, une évolution exclusivement technologique de COLCHIC permet de passer d'un système décentralisé vers une solution centralisée. Sous COLCHIC 2, chaque LIC possédait, sur un serveur local, sa base propre et une réplication de la base nationale. Les échanges de données entre tous ces composants se faisaient périodiquement. Le système centralisé, développé sous COLCHIC 3, consiste à héberger à l'INRS, un seul serveur accessible via Internet pour tous les utilisateurs autorisés. Aucune modification liée à la codification de l'information n'a été entreprise lors de l'évolution de COLCHIC 2 vers COLCHIC 3.

III.1.2. SCOLA

III.1.2.1. Généralités

A la demande de la Direction Générale du Travail (DGT), l'INRS développe en 2006 le système de collecte des informations pour les laboratoires accrédités (SCOLA). Cette base de données, destinée à centraliser les résultats de mesures d'exposition aux agents chimiques disposant d'une valeur limite d'exposition professionnelle (VLEP) réglementaire, sera renseignée par les organismes accrédités pour la réalisation de ces contrôles d'exposition. La structure développée pour SCOLA repose sur celle de COLCHIC. Le système de codification est ainsi repris en l'état pour certaines variables ou adapté aux spécificités de SCOLA pour d'autres. La mise en production de SCOLA a été réalisée début juin 2007, suite à la parution de l'arrêté du 4 mai 2007 relatif à la mesure de la concentration en fibres d'amiante sur les lieux de travail et aux conditions d'accréditation des laboratoires.

Dans un premier temps, seuls les résultats des contrôles d'exposition à l'amiante, effectués sous accréditation, devaient systématiquement être renseignés dans la BDEP. Sur la base du volontariat, il est rendu possible de 2007 à 2009 pour les organismes accrédités, d'enregistrer les mesures effectuées sous accréditation pour une vingtaine

d'agents chimiques. Dans un deuxième temps, le renforcement de la réglementation, avec la mise en place du décret n°2009-1570 du 15 décembre 2009 et de son arrêté d'application du même jour, a étendu la saisie dans la base de données à 134 substances avec une VLEP réglementaire (articles R4412-149 et R4412-150 du Code du travail). Le décret amiante n°2012-639 du 4 mai 2012 a instauré les contrôles de l'empoussièrement en amiante par microscopie électronique à transmission analytique (META) en remplacement de la technique par microscopie optique à contraste de phase (MOCP).

Sur la période 2007-2015, SCOLA compte 33 agents chimiques avec, pour chacun, plus de 1 000 résultats validés ou archivés.

III.1.2.2. Exploitation des données de SCOLA

Un rapport d'activité produit par l'INRS, présentant les niveaux d'exposition enregistrés dans SCOLA par agent chimique, est transmis annuellement à la DGT. Dans le cadre du décret amiante de 2012, un rapport annuel complémentaire et spécifique portant sur les niveaux *a priori* d'empoussièrement par couple matériau technique a été rédigé jusqu'en 2014. Depuis, ce rapport a été remplacé par une application développée par l'INRS et nommée Scol@miante, accessible sur <http://scolamiante.inrs.fr/>.

III.1.2.3. Les évolutions de SCOLA

Les évolutions de SCOLA suivent principalement celles de la réglementation. Ainsi, toute modification du Code du travail impactant les contrôles réglementaires peut nécessiter un aménagement de SCOLA. Par exemple, la mise en application de l'arrêté du 15 décembre 2009 et du décret 2012-639 du 4 mai 2012 a eu pour conséquences la mise à jour de tables de référence servant à la codification des mesures et le développement de nouvelles fonctionnalités de SCOLA : saisie des variables « Type de lieu de travail », « Ventilation globale », « Protection collective » et « Protection individuelle » devenue obligatoire fin 2009 et création de variables spécifiques telles que « Matériau », « Technique » pour les mesurages d'amiante en 2012.

III.2. PRÉPARATION DES DONNÉES

Préalablement à toute analyse des informations issues d'une BDEP, un travail préliminaire de sélection des données est à mettre en œuvre. En fonction des objectifs et des besoins, plusieurs étapes peuvent être nécessaires.

III.2.1. CRÉATION DES JEUX DE DONNÉES POUR LES ANALYSES

III.2.1.1. Extraction et critères de choix pour les différentes analyses

Afin de répondre aux objectifs de l'étude, il nous a semblé pertinent d'extraire la totalité des données enregistrées dans COLCHIC et SCOLA. Au démarrage de l'étude, une première extraction, de 1987 à 2012 pour COLCHIC et de 2007 à 2012 pour SCOLA a permis la réalisation de leurs analyses descriptive et comparative. Puis, pour les analyses séparées par modélisation des deux BDEP, les données ont été complétées avec les enregistrements des années 2013, 2014 et 2015. SCOLA a ainsi été analysée sur 2007-2015. Afin de tenir compte des évolutions de la BDEP, deux périodes distinctes ont été modélisées pour COLCHIC : 1987-2015, portant sur la globalité des mesures et 2002-2015, incluant les nouvelles variables codées à partir de 2002. Enfin, la modélisation comparative de COLCHIC et SCOLA s'est portée sur la période commune aux deux BDEP, 2007-2015.

Après chaque extraction des BDEP, les données répondant aux critères suivants ont été conservées :

- ▶ Les mesures quantitatives d'air des lieux de travail réalisées dans la zone respiratoire des travailleurs (mesures individuelles) et à point fixe (mesures d'ambiance) ;
- ▶ Les concentrations exprimées en unités d'air ;
- ▶ Les durées de prélèvement comprises entre 1 et 720 minutes.

III.2.1.2. Variables à l'étude

L'un des objectifs de ce travail est d'étudier les liens possibles entre les niveaux

d'exposition dans COLCHIC et SCOLA et certaines informations accompagnant les mesures, « déterminants » potentiels de l'exposition. COLCHIC et SCOLA contiennent une grande quantité d'informations contextuelles, incluant notamment des informations de type administratif, non pertinentes pour ce travail. Le Tableau III.1 présente les informations documentées dans les deux BDEP qui ont été considérées dans nos analyses.

► **L'évolution temporelle des variables dans les BDEP**

La robustesse de l'exploitation d'une BDEP repose, entre autres, sur la stabilité de la codification de l'information dans le temps. C'est-à-dire que les évolutions, liées à la création de nouvelles catégories au sein d'une variable, à leurs modifications ou même à leurs suppressions, doivent être raisonnées, maîtrisées et tracées par les gestionnaires de la BDEP. Ces changements, pouvant être liés à des retours d'expérience, sont néanmoins nécessaires pour suivre les évolutions techniques et organisationnelles du monde industriel et ainsi garantir une codification des informations au plus près de la réalité du terrain. COLCHIC et SCOLA n'ont pas échappé à cette règle. C'est pourquoi, au cours du déploiement de nouvelles versions de COLCHIC et SCOLA, la création ou la modification de plusieurs variables ont été nécessaires. Par exemple, comme nous l'avons déjà vu, les métiers, le type de procédé et la fréquence d'exposition ont été créés dans COLCHIC 2 avec comme objectif de renforcer la connaissance du poste de travail et de son environnement. Ces changements ne sont pas sans conséquence sur l'exploitation des BDEP. Par exemple, pour COLCHIC, ces 3 variables sont exploitables uniquement après 2002.

► **Création de nouvelles variables**

Parmi les informations enregistrées dans une BDEP, certaines peuvent nécessiter d'être transformées, modifiées ou associées pour pouvoir les exploiter. C'est ce que nous avons été amenés à effectuer dans 4 situations : 1- création d'une variable région à partir du code postal de la ville où est localisée l'entreprise investiguée 2- enregistrement de la durée de chaque prélèvement dans COLCHIC et SCOLA. Cette variable a été utilisée pour créer une variable dichotomique « Type d'échantillonnage » caractérisée comme « Courte

durée » si la durée est inférieure à 60 minutes et comme « Longue durée » si la durée est supérieure ou égale à 60 minutes 3- création d'une variable « Durée stratifiée » comptant 4 catégories de durée de prélèvement (d1 : 1 à 60 minutes, d2 : 61 à 120 minutes, d3 : 121 à 240 minutes et d4 : 241 à 720 minutes) 4- identification de chaque intervention à partir d'une variable formée par le triptyque « Année », « Laboratoire » et « Numéro de dossier ».

► **Regroupements de catégories**

Certaines catégories au sein d'une même variable (par exemple, « Absence de ventilation mécanique » et « Dispositif hors fonction » pour la variable « Ventilation globale ») ont été regroupées lorsqu'elles contenaient trop peu d'observations pour faciliter l'analyse.

► **Corrélation entre les variables**

Le degré de corrélation entre les déterminants potentiels de l'exposition, identifiés préalablement pour étudier leur lien avec le niveau d'exposition mesuré, a été évalué au moyen de différents indices afin d'éviter des problématiques de multicolinéarité.

- Pour A et B, deux variables continues, le coefficient de corrélation linéaire de Pearson a été calculé ;
- Pour A et B, deux variables nominales, l'association entre A et B est estimée à partir du coefficient de Cramer's V ;
- Pour A, une variable continue et B, une variable nominale, le coefficient de détermination d'une analyse de la variance (ANOVA) a permis de quantifier l'association entre A et B.

Deux variables sont considérées comme trop fortement dépendantes si le coefficient calculé est supérieur ou égal à 0,7 [Sarazin et al., 2016]. Si deux variables sont corrélées au-delà de ce seuil, alors l'une d'entre elles ne sera pas incluse dans les analyses statistiques.

III.2.1.AGGLOMÉRATION DES VARIABLES D'ACTIVITÉ ÉCONOMIQUE

Les variables décrivant l'« Activité économique » (métier, secteur d'activité, tâche) comptent plusieurs centaines de catégories, organisées en niveaux hiérarchiques. Par exemple, pour la classification ROME qui permet de décrire le métier, il existe 4 niveaux de codification : 8 digits pour caractériser le niveau le plus fin puis 5 et 3 digits pour les niveaux intermédiaires et enfin 2 digits pour le niveau le plus grossier. Afin d'assurer un nombre suffisant de données par catégorie, les catégories possédant un faible nombre de données ont été agglomérées. Cette approche est employée fréquemment pour la gestion de ce type de variables [Lavoué et al., 2008 ; Lee et al., 2015 ; Sarazin et al., 2016]. Pour cela, un « critère », défini comme le nombre minimal de données appartenant à une catégorie, est calculé comme suit :

$$Critère = n_{med.prelvt} \times x$$

$n_{med.prelvt}$: nombre médian de prélèvements par entreprise pour un agent chimique

x : facteur multiplicatif

Le mode de calcul du critère permet d'ajuster sa valeur aux différentes situations enregistrées dans les BDEP. Il permet ainsi de s'assurer, pour un agent chimique donné, que l'on a probablement plusieurs établissements représentés dans chaque catégorie. Ainsi, si le nombre de données dans une catégorie est inférieur au critère, alors ces données seront agglomérées à un niveau hiérarchique supérieur.

Tableau III.1 Principales variables documentées dans COLCHIC et SCOLA

Variable	^a Cat. information	Type de données	Définition
COLCHIC et SCOLA			
Propriétaire	E	Catégorie	Laboratoire ou organisme propriétaire du dossier
Année	E	Date	Année de réception de la demande d'intervention (COLCHIC) ou de réalisation des prélèvements (SCOLA)
N° Dossier	E	Catégorie	N° du dossier de la demande d'intervention
SIRET	E	Catégorie	Identifiant de l'établissement d'une entreprise en tant qu'unité géographiquement localisée
Etablissement – Ville	E	Catégorie	Coordonnées de l'établissement (adresse, code postal)
NAF	E	Catégorie	Code de la nomenclature des Activités Françaises (NAF rév 02, 2008)
Effectif	E	Catégorie	Tranche d'effectif salarié de l'établissement
Préleveur	CP	Catégorie	Désignation de la personne (COLCHIC) ou de l'organisme (SCOLA) qui a assuré la responsabilité de la campagne de prélèvements dans l'établissement
Date prélèvement	CP	Date	Date de réalisation des prélèvements dans l'établissement
Référence / N° de Prélèvement	CP	Catégorie	Référence du prélèvement donnée par le préleveur
Méthode de prélèvement	CP	Catégorie	Permet de caractériser la technique mise en œuvre pour réaliser le prélèvement
Type support	CP	Catégorie	Permet de caractériser le dispositif de prélèvement (nature et dimensions)
Débit (L/min)	CP	Continue	Débits mesurés au début et à la fin d'un prélèvement actif
Objectif**	CP	Catégorie	Type de prélèvement effectué (ambiance ou individuel ; quantitatif ou qualitatif)
Représentativité**	CP	Catégorie	Représentativité du prélèvement indiquée par le préleveur compte tenu de sa connaissance des conditions d'exposition
Type de lieu de travail**	Ex	Catégorie	Description sommaire du type de lieu de travail où a été réalisé le prélèvement
Ventilation globale**	Ex	Catégorie	Description sommaire de la ventilation générale du lieu de travail
Protection collective**	Ex	Catégorie	Description des moyens de protection collective mis en œuvre
Tâche	CP	Catégorie	Description générale du type de tâche réalisée pendant des prélèvements
Métier*	CP	Catégorie	Identification du type de métier exercé par le salarié (Répertoire Opérationnel des Métiers et des Emplois du pôle emploi (ROME version 1.2 – 1999))

Variable	^a Cat. information	Type de données	Définition
Protection individuelle	CP	Catégorie	Description des moyens de protection individuelle mis en œuvre
Durée (min)	CP	Continue	Durée en minutes du prélèvement
Observations prélèvement	CP	Texte	Observations relevées au cours du prélèvement
Laboratoire	CA	Catégorie	Laboratoire en charge de l'analyse demandée
Agent Chimique	CA	Catégorie	Agent chimique mesuré lors de l'analyse
Technique d'analyse	CA	Catégorie	Technique d'analyse utilisée pour quantifier l'agent chimique
COLCHIC seule			
Risque	E	Catégorie	N° de risque de l'établissement
Origine de demande	E	Catégorie	Indique l'origine déterminante de la demande d'intervention
Motif de la demande	E	Catégorie	Indique le motif déterminant de la demande d'intervention
Lieu de prélèvement	E	Texte	Localisation générale du lieu où sont réalisés les prélèvements
Lieu de travail	E	Texte	Nom du lieu de travail (atelier, local...) où a été réalisé le prélèvement
Activité production estimée*	Ex	Catégorie	Description de l'activité de production du lieu de travail où sont réalisés les prélèvements
Température ambiante	Ex	Catégorie	Intervalle de température ambiante du lieu de travail pendant les prélèvements
Type de procédé*	Ex	Catégorie	Description du type de procédé générant les agents chimiques à l'origine de l'exposition
Fréquence d'exposition*	Ex	Catégorie	Caractérisation de la fréquence d'exposition à la catégorie de produit ou du procédé à l'origine de l'exposition
CAPPROEX*	Ex	Catégorie	Catégorie de produit ou de procédé à l'origine de l'exposition professionnelle
T° utilisation*	Ex	Texte	Température d'utilisation du produit à l'origine de l'exposition

^aE, information administrative ; CP, conditions de prélèvement ; Ex, conditions d'exposition ; CA, conditions analytiques

* ajouté en 2002 dans COLCHIC

** modifié en 2002 dans COLCHIC

III.2.2. TRAITEMENT DES NON DÉTECTÉS

Lors de la réalisation d'une campagne de mesures dans l'air des lieux de travail, il arrive que la concentration d'un agent chimique ne puisse être quantifiée [Chapitre IV] : soit parce que l'agent chimique n'est pas présent dans la zone de travail où la mesure a lieu, soit parce que la concentration est trop faible pour être quantifiée par le système analytique déployé. Ces résultats sont appelés « Non-Détecté ». Au fil du temps, la codification utilisée dans COLCHIC pour caractériser ce type de résultat a évolué. Pendant plusieurs années, il était courant de coder les « non détectés » avec l'un de ces deux acronymes : « DE » ou « ND », respectivement « Détecté » et « Non Détecté ». Dans ce cas, aucune information numérique n'a été enregistrée rendant rétrospectivement ces mesures difficiles à exploiter. Aujourd'hui, l'emploi de « DE » ou « ND » n'est plus possible. Lorsqu'un laboratoire ne parvient pas à quantifier une concentration, l'utilisateur saisit dans COLCHIC/SCOLA le symbole « < » ou « ≤ » en accompagnement d'une valeur numérique. Cette dernière est la limite de quantification (LoQ) exprimée sous la forme d'une concentration. Afin de conserver l'ensemble de ces résultats « non détectés », notamment ceux codés « ND » et « DE », il a été nécessaire de les exprimer sous la forme d'une concentration. Pour cela, la gestion des « DE »/« ND » s'est faite en définissant une LoQ empirique. 3 périodes (p1 : 1987-1995, p2 : 1996-2001 et p3 : 2002-2015) et 4 intervalles de durée de prélèvement (d1 : 1 à 60 minutes, d2 : 61 à 120 minutes, d3 : 121 à 240 minutes et d4 : 241 à 720 minutes) ont été définis [Raaschou-Nielsen et al. 2010]. Pour chaque agent chimique appartenant à chacun de ces couples période/durée (par exemple, p1/d1, 1987-1995 pour les prélèvements d'une durée comprise entre 1 et 60 minutes), la concentration minimale mesurée a été utilisée en remplacement des « DE » et des « ND ».

III.3. MODÉLISATION STATISTIQUE POUR CHAQUE AGENT CHIMIQUE

Afin d'identifier dans un premier temps, les variables associées aux niveaux d'exposition mesurés dans COLCHIC et SCOLA, puis dans un deuxième temps, les différences systématiques des niveaux mesurés dans chacune des 2 BDEP, nous avons utilisé

plusieurs modèles statistiques.

Cette section présente les types de modèles utilisés, puis les étapes qui ont permis de définir leur structure et enfin l'exploitation des résultats.

III.3.1. GÉNÉRALITÉS SUR LES MODÈLES LINÉAIRES MIXTES

Burstyn et al. ont proposé, en 1999, une revue des recherches rapportant des résultats d'études sur les déterminants de l'exposition et montrant l'association entre l'exposition et des facteurs associés aux niveaux d'exposition. La majorité de ces études se sont appuyées sur des modèles linéaires simples et des modèles de régression linéaires multiples. Ces approches sont maintenant utilisées dans la majorité des études de ce type, en particulier pour les identifications de variables importantes des BDEP [Teschke et al., 1999 ; Lavoué et al., 2005 ; Zuur e tal., 2009 ; Clerc et al., 2014 ; Lee et al., 2015 ; Sarazin et al., 2016].

Un modèle linéaire simple est un modèle ayant pour but d'expliquer une variable dépendante (la concentration dans notre étude) en fonction de variables indépendantes, également appelées prédictors. En pratique, il permet d'établir le lien entre des prédictors et la variable réponse. Généralement, il est utilisé pour répondre à deux objectifs principaux : 1- prédire une valeur moyenne de Y en fonction de X [Lavoué et al., 2008 ; Clerc et al., 2015a] 2- mesurer l'association entre X et Y [Lavoué et al., 2006 ; Sauvé et al., 2012, Sarazin et al., 2016]. Dans notre situation, nous souhaitons expliquer les concentrations enregistrées dans COLCHIC et SCOLA (variable quantitative) en fonction de leurs variables descriptives.

Pour les modèles linéaires à effets mixtes, également appelés modèles hiérarchiques, l'appellation « mixte » signifie qu'en plus des prédictors appelés « effets fixes », une structure de regroupement des données peut être modélisée (par exemple, des mesures répétées pour un travailleur [Burstyn et al. 1997], ou des mesures prises dans un même établissement ou durant une même intervention [Kromhout et al. 1994]). La non-prise en compte des phénomènes de corrélation à l'intérieur d'une structure de groupe peut causer la sous-estimation de l'incertitude et biaiser les estimations [Lavoué et al., 2005].

Une variable spécifiée comme identifiant une structure de regroupement est appelée « aléatoire » et nous parlerons d'effet aléatoire par opposition à effet fixe.

III.3.2. MODÈLE TOBIT

III.3.2.1. Limites des modèles linéaires pour les données censurées

Comme nous l'avons préalablement expliqué, lors de la réalisation de campagnes de prélèvements, la quantification de l'agent recherché n'est pas systématique. Fréquemment, les résultats sont notés « Non Détecté », qualificatif équivalent à celui de « Censuré » dans le domaine de l'épidémiologie. L'une des limites d'un modèle linéaire est qu'il ne permet pas l'inclusion dans l'analyse de données « censurées ». Dans de telles circonstances, l'utilisation de modèles mieux adaptés est alors recommandée.

Pour permettre de décrire la relation entre une variable dépendante censurée et une variable indépendante (explicative ou variable exogène), l'emploi d'un modèle de régression censuré tel que le modèle Tobit est généralement recommandé [Tobin, 1958 ; Lubin et al., 2004]. Néanmoins, en fonction de la proportion de valeurs censurées, plusieurs approches plus simples restent envisageables. Par exemple, lorsque la proportion de mesures censurées est relativement faible (5 à 10 %), l'utilisation de modèles linéaires classiques en remplaçant simplement x par $x/2$ [Kauffer et Vincent, 2006, Clerc et al., 2014] ou $x/\sqrt{2}$ [Hornung et Reed, 1990, Agostini et al., 2010] reste possible. Dans la situation où la proportion de mesures censurées est importante (> 30 %) l'emploi du modèle Tobit [Tobin, 1958 ; Lubin et al., 2004 ; Lavoué et al., 2010 ; Sauvé et al., 2012] ou de l'imputation multiple [Lubin et al., 2004 ; Peters et al., 2011] est alors préconisé.

Au regard de l'analyse descriptive sur la période 2007 à 2012 pour les 17 agents chimiques communs à COLCHIC et SCOLA présentée chapitre IV, les proportions médianes de mesures individuelles « longue durée » inférieures à la LoQ sont respectivement de 15 % [1 % pour l'ammoniac anhydre - 86 % pour la cristobalite] et de 58 % [13 % pour l'ammoniac anhydre - 92 % pour la cristobalite]. Il en ressort, d'une manière générale, que la proportion importante de mesures censurées par agent chimique dans COLCHIC et

SCOLA confirme notre intérêt à utiliser le modèle Tobit. Nous avons choisi un seuil maximum de 60 % de « Non Détecté » comme critère de sélection des agents chimiques.

III.3.2.2. Principe des modèles TOBIT

Les modèles Tobit sont qualifiés de modèles de survie. Cela signifie que, dans le domaine de la biologie par exemple, ce type de modèle est employé pour représenter le temps de survie des patients en fonction de différents paramètres : les échantillons sont considérés comme censurés dès lors que le patient est en vie à la dernière date d'observation de l'échantillon ; on parle alors d'observations censurées à droite ($>x$). Les modèles Tobit sont des modèles de survie adaptés à des observations censurées à droite. Ils sont employés pour évaluer par exemple, l'évolution des niveaux de concentration dans le temps en fonction de descripteurs [Lavoué et al., 2010 ; Sauvé et al., 2015]. Ce type de modèle permet de prendre en compte l'ensemble des situations enregistrées en y incluant un effet aléatoire.

Les données d'exposition en milieu du travail suivent généralement la distribution normale [Guide méthodologique MétroPol - Interprétation statistique des résultats de mesure, 2015]. Une transformation logarithmique, pour obtenir une distribution normale des observations faisant partie des hypothèses des modèles linéaires, a donc été réalisée [Gomez, 1997 ; Hornung et Reed, 1990 ; Melville et Lippmann, 2001].

III.3.3. STRATÉGIE DE MODÉLISATION

Tout effort de modélisation statistique nécessite la sélection des prédicteurs qui seront inclus dans le modèle final. Ce processus, qui peut aller de l'inclusion de toutes les variables disponibles à une sélection restreinte basée sur la comparaison de plusieurs modèles plausibles, ne fait pas l'objet de consensus et peut varier selon les objectifs de l'étude. Ce travail doctoral comportait la contrainte supplémentaire d'une étape post-modélisation de synthèse des résultats à travers plusieurs dizaines d'agents chimiques.

Nous avons donc opté pour une approche d'inclusion de toutes les variables jugées d'intérêt (Tableau III.1). Cette liste de variables varie en fonction de la période considérée

(par exemple, COLCHIC 1987-2015 / COLCHIC 2002-2015), mais pour une même période, la même liste a été utilisée pour tous les agents chimiques.

Bien qu'une telle approche soit vulnérable à l'observation de coefficients « significatifs » malgré l'absence d'une réelle association, il est peu probable qu'une association due à la chance apparaisse pour de multiples agents. L'étape de Méta analyse (décrite plus bas) devrait donc considérablement atténuer la problématique.

Pour 3 structures particulières nommées « Aléatoire », « Nature de l'activité économique » et « Tendances temporelles et durée de prélèvement », plusieurs options de modélisation étaient disponibles. A ce jour, les logiciels permettant de calculer les modèles Tobit sont limités à un seul effet aléatoire. Le choix s'est fait en se référant aux données issues de la littérature scientifique. Pour les deux autres structures, il a été effectué en déterminant en moyenne à travers tous les agents, celle qui semblait la meilleure selon deux critères différents : l'Akaike Information Criterion (AIC) et le Bayesian Information Criterion (BIC).

L'AIC, proposé par Hirotugu Akaike en 1973, repose sur le nombre de paramètres à estimer (k) et le maximum de la fonction de vraisemblance du modèle (V). Il s'écrit :

$$AIC = 2k - 2 \ln(V)$$

Le BIC, proposé par Gideon Schwarz en 1978, est similaire à l'AIC. Dans son calcul, il prend en plus en compte, la taille de l'échantillon. Il s'écrit :

$$BIC = 2k \times \ln(N) - 2 \ln(V)$$

Le choix de l'AIC et du BIC comme critères de qualité repose sur le fait que dans un modèle, l'ajout d'une variable a toujours pour conséquence l'amélioration de sa fonction de vraisemblance. Ainsi, plus un modèle comptera de variables, plus sa vraisemblance sera renforcée. Dans ce contexte, il est difficile sur ce seul critère de juger de la performance du modèle. Afin de pondérer cette approche, l'emploi de l'AIC et du BIC comme critères de qualité, permet de pénaliser les modèles avec un grand nombre de paramètres, la pénalité étant plus forte et dépendante de la taille d'échantillon dans le cas du BIC. Pour les deux critères, les modèles avec les valeurs les plus basses sont retenus comme plus performants.

III.3.3.1. La structure « Aléatoire »

Plusieurs études comme celles de Lavoué et al. [2005] et Sarazin et al. [2016] ont mis en évidence l'intérêt d'inclure une structure aléatoire. En effet, ne pas tenir compte de cette structure de corrélation, c'est-à-dire utiliser un simple modèle linéaire, en plus de ne pas permettre son évaluation quantitative, engendrerait une sous-estimation de l'incertitude, voire des biais dans les coefficients. Par exemple, dans le domaine de l'épidémiologie, ces structures sont utilisées pour tenir compte d'unités comme les hôpitaux ou les écoles à l'intérieur desquelles on soupçonne une similarité de l'état observé. En hygiène industrielle, ces modèles ont été utilisés pour évaluer les groupes tels que les travailleurs, les départements, les établissements ou encore les interventions. Lors des analyses de COLCHIC et SCOLA, ces deux dernières structures ont été explorées. La première structure correspond à l'établissement dans lequel la ou les campagnes de mesures ont été effectuées. La seconde structure est l'intervention à un instant donné dans un établissement : par exemple, deux campagnes de prélèvements deux années de suite dans un même établissement donnent lieu à deux interventions. Cette dernière structure utilisée par Lavoué et al. [2011] est celle que nous avons retenue pour l'ensemble des analyses.

III.3.3.2. Modélisation de la structure « Nature de l'activité économique »

Plusieurs variables codées dans COLCHIC et SCOLA permettent de décrire l'« Activité économique » de l'entreprise ou du travailleur : la nomenclature des activités françaises (NAF), le métier ou encore la tâche. Lorsque la modélisation intègre les données antérieures à cette date, l'enregistrement du métier, à partir de 2002 dans COLCHIC, a pour conséquence un choix réduit pour définir la structure « Activité économique ». Par exemple, pour la période 1987 à 2015, le choix s'est fait entre les variables NAF et tâche. Pour les autres périodes démarrant à partir de 2002, la variable métier a pu être ajoutée. Cette structure « Activité économique », dans le cadre de nos travaux, est utilisée comme une variable de confusion. Elle permet ainsi de prendre en considération la variabilité des concentrations liée à l'activité. Elle n'a pas pour vocation d'être exploitée. Le choix s'est porté sur celle étant la plus prédictive (c'est-à-dire avec les meilleurs AIC et BIC) pour

chacune des périodes analysées.

Pour le choix de cette structure, nous avons comparé des modèles comprenant les variables jugées d'intérêt, la structure « Tendances temporelles et durée de prélèvement » la plus complète et la structure aléatoire « Intervention » préalablement déterminée auxquelles nous avons ajouté tour à tour une variable de la structure « Activité économique ».

III.3.3.3. Modélisation de la structure « Tendances temporelles et durée de prélèvement »

Les tendances temporelles et la durée de prélèvement jouent un rôle essentiel dans l'estimation des niveaux de concentration [Symanski et al., 1998 ; Kolstad et al., 2005 ; Lavoué et al., 2005 et 2010 ; Flanagan et al., 2006 ; Raaschou-Nielsen et al., 2010, Sauvé et al., 2012]. Récemment, à travers les travaux de Clerc et Vincent [2014], il a été montré que plus la durée d'échantillonnage est courte plus la concentration mesurée est élevée.

Cette structure « Tendances temporelles et durée de prélèvement » est caractérisée dans nos travaux à partir des deux variables continues « Année » et « Durée de prélèvement », de la variable dichotomique « Type d'échantillon » (courte durée – longue durée) et de la variable de type catégorie « Durée stratifiée ». Les travaux de Lavoué et al. [2005] laissent présager que les variables de cette structure interagissent. De plus, Lavoué et al. [2010] ont mis en évidence, entre autres, des interactions entre l'année et le type d'échantillon noté « Année : Type_échantillon » dans nos structures. Pour tenir compte de possibles interactions entre ces variables, neuf combinaisons de structures avec et sans interaction ont été testées avant de choisir celle retenue pour le modèle définitif :

- (1) Année seule
~ Année
- (2) Durée stratifiée seule
~ Durée_stratifiée
- (3) Année en interaction avec durée stratifiée
~ Année + Durée_stratifiée + Année:Durée_stratifiée
- (4) Durée de prélèvement seule

- ~ Durée_prélèvement
- (5) Type d'échantillon seul
 - ~ Type_échantillon
- (6) Année seule puis durée de prélèvement seule et enfin type d'échantillon seul
 - ~ Année + Durée_prélèvement + Type_échantillon
- (7) Durée de prélèvement seule, puis année en interaction avec type d'échantillon
 - ~ Année + Durée_prélèvement + Type_échantillon + Année:Type_échantillon
- (8) Année seule, puis durée de mesure en interaction avec type d'échantillon
 - ~ Année + Durée_prélèvement + Type_échantillon +
Durée_prélèvement:Type_échantillon
- (9) Année en interaction avec type d'échantillon puis, durée de prélèvement en interaction avec type d'échantillon
 - ~ Année + Durée_prélèvement + Type_échantillon + Année:Type_échantillon +
Durée_prélèvement:Type_échantillon

Pour le choix de cette structure, nous avons comparé des modèles comprenant les variables jugées d'intérêt, la structure « Activité économique » choisie, la structure aléatoire « Intervention » préalablement déterminée auxquelles nous avons ajouté tour à tour une combinaison de la structure « Tendances temporelles et durée de prélèvement ».

Enfin, pour une période définie, la structure de modèle final, identique pour tous les agents, incluait donc toutes les variables d'intérêt en plus d'une combinaison des structures décrites ci-dessus. En termes techniques, le choix de chacune des structures a été effectué dans l'ordre suivant [Zuur et al., 2009] : 1- choix de la structure aléatoire 2- choix de la structure « Activité économique » 3- choix de la structure « Tendances temporelles et durée de prélèvement ». Les comparaisons entre différentes structures d'effet aléatoire (1-) employaient la méthode d'estimation « restricted maximum likelihood », également utilisée pour estimer les coefficients finaux. Les comparaisons entre différentes structures d'effet fixe (2- et 3-) employaient la méthode d'estimation « maximum likelihood ».

III.3.4. MESURE DE LA QUALITÉ DU MODÈLE

La qualité du modèle revient à mesurer la proportion de la variabilité des niveaux d'exposition expliquée par les effets fixes. Nous avons donc calculé le coefficient de détermination (R^2) pour nous faire une idée des performances du modèle [Burstyn et Techke, 1999 ; Lavoué et al., 2005 ; Sauvé et al., 2015 ; Sarazin et al., 2016].

Plus le coefficient de détermination est proche de 1, meilleur est le modèle. C'est-à-dire que les effets fixes expliquent une importante proportion de la variabilité des niveaux d'exposition. À l'inverse, plus le coefficient est proche de 0, moins bon est le modèle.

Les R^2 ont été calculés à travers tous les agents chimiques exploités, stratifiés par période et par type de prélèvement. La valeur médiane par strate a été définie afin d'apprécier la qualité générale du modèle.

III.3.5. INTERPRÉTATION DES COEFFICIENTS ISSUS DE LA MODÉLISATION

Dans notre étude, un coefficient peut se définir comme étant la différence entre la moyenne arithmétique des logarithmes de l'exposition pour la catégorie d'intérêt et celle de la catégorie de l'intercepte.

Pour les variables de type catégorie, les coefficients des modèles sont exploités en utilisant un indice d'exposition relative (RIE). Ce dernier est défini comme étant l'exponentielle de la différence entre 2 coefficients estimés [Lavoué et al., 2005 ; Sauvé et al., 2012 ; Sarazin et al., 2016]. Il s'écrit :

$$RIE_{Cat A} = e^{(Coeff_{Cat A} - Coeff_{Cat Ref})}$$

$RIE_{Cat A}$: Indice d'exposition relative de la catégorie A de la variable considérée

$Coeff_{Cat A}$: Coefficient estimé de la catégorie A de la variable considérée

$Coeff_{Cat Ref}$: Coefficient estimé correspondant à la catégorie de référence de la variable considérée.

Le $Coeff_{Cat Ref}$ est égal à 0 lorsque la catégorie de référence est incluse dans l'intercepte. Afin d'illustrer son interprétation, prenons par exemple, un RIE de 0,50 d'une catégorie A d'une variable. Cela signifie qu'en moyenne, l'exposition est deux fois plus faible dans la

situation de la catégorie A comparée à celle de la catégorie de référence.

Pour les variables de type continu, les exponentielles des coefficients des modèles correspondent à des pentes et sont directement exploitées. Par exemple, l'exponentielle d'un coefficient de 0,95 pour un taux de réduction annuelle de l'exposition signifierait qu'annuellement le niveau des concentrations décroît de 5 %. La durée de prélèvement est ainsi exploitée dans nos analyses à travers un taux de réduction de l'exposition, provoqué par une augmentation de 30 minutes de la durée de prélèvement et stratifié par « Courte durée » et « Longue durée ». Pour la tendance temporelle, il s'agit d'un taux de réduction annuelle également stratifié par type d'échantillon. Pour le calcul de ces taux, lorsque les mesures de « Courte durée » ont été écartées du jeu de données, compte tenu de leur proportion inférieure à 20 %, le taux de réduction de l'exposition pour l'agent chimique concerné n'a été calculé que pour les prélèvements « Longue durée ».

Enfin, dans le but de comparer les niveaux d'exposition entre les prélèvements « Courte durée » et « Longue durée », nous avons calculé pour la première et la dernière année du jeu de données, le ratio entre les concentrations mesurées sur une durée de 240 minutes et celles mesurées sur une durée de 30 minutes.

Pour chaque agent chimique, on obtient ainsi la même liste de coefficients et d'erreurs standards, ce qui se prête parfaitement à la procédure de Méta analyse.

III.4. MÉTA ANALYSE

L'exploitation des coefficients estimés par les modèles pour plusieurs dizaines d'agents chimiques nécessite une approche permettant de synthétiser l'ensemble de ces résultats. Pour cela, la méthode Méta analyse, fréquemment utilisée en épidémiologie, a été employée.

III.4.1. PRINCIPE GÉNÉRAL

Glass [1976] définit la Méta analyse comme « une revue systématique de la littérature soutenue par des méthodes statistiques où le but est d'agréger et de comparer les

résultats de plusieurs études connexes ». Cette approche, dont l'objectif est de synthétiser un ensemble de résultats, est utilisée aujourd'hui dans des domaines variés tels que la médecine, l'épidémiologie, l'expologie... [Houwelingen et al., 2002 ; Borenstein et al., 2010 ; Viechtbauer, 2010 ; Sarazin et al., 2016]. Dans des milieux où l'information se fait rare, la Méta analyse permet d'améliorer la connaissance par la combinaison d'études indépendantes et d'en sortir des conclusions globales.

Ces conclusions résultent de « pondérations » des résultats des études et sont basées sur le calcul d'un « effet global » assimilé à la tendance générale.

Afin de combiner les effets de plusieurs études indépendantes pour estimer un effet global, certains paramètres doivent être pris en compte [Littel et al., 2008]. En effet, chaque étude ne contribue pas de manière égale à la détermination de l'effet global. Son poids peut varier en fonction de l'intervalle de confiance, de ses résultats, mais aussi de la taille d'échantillon de l'étude. Ainsi, plus les résultats d'une étude sont robustes (intervalle de confiance étroit) ou plus l'échantillon d'une étude est important (bonne représentativité), plus l'étude aura de poids dans la Méta analyse.

L'estimation de l'effet global dépend aussi du type de modèle utilisé dans la Méta analyse [Borenstein et al., 2010] : modèle à effets fixes, modèle à effets aléatoires ou modèle à effets mixtes. Dans le cas des modèles à effets fixes, seule la variation intra-étude est prise en compte. La variation inter-étude est assimilée uniquement à l'erreur sur l'échantillonnage. Dans le cas des modèles à effets aléatoires, c'est différent puisque l'on considère que l'effet peut aussi varier d'une étude à une autre. Enfin, les modèles à effets mixtes permettent de combiner les avantages des deux autres méthodes. Ainsi, avec ce type de modèle, les études sont assimilées à des variables modératrices de l'effet moyen.

Le résultat d'une Méta analyse est exprimé sous la forme d'un coefficient global – ou Méta coefficient – qui représente l'effet moyen à travers toutes les études. Il est généralement associé à un intervalle de confiance. Des Forest plots sont aussi utilisés pour illustrer, sur un même graphique, les résultats des études et de la Méta analyse et identifier la variabilité entre les études.

III.4.2. APPLICATION À NOTRE ÉTUDE

Dans le cadre de nos travaux, nous avons suivi la même approche : les études sont remplacées par les agents chimiques. Pour chaque catégorie de variables, un Méta coefficient et son intervalle de confiance à 95 % sont calculés à travers tous les agents chimiques [Houwelingen et al., 2002 ; Olsson et al., 2010 ; Sarazin et al., 2016]. Ce Méta coefficient est ensuite exprimé sous la forme de RIE pour les variables de type catégorie ou sous la forme de pente moyenne à tous les agents chimiques pour les variables continues. Les coefficients de la structure « Activité » ne sont quant à eux pas exploités, puisque d'une part son rôle est d'être utilisée comme une variable de confusion et d'autre part les catégories ne sont pas les mêmes à travers les agents chimiques.

La représentation graphique sous forme de « Forest plot » permet d'illustrer la variabilité interagents chimiques.

Enfin, la combinaison des résultats par Meta analyse a permis de répondre à deux objectifs : 1- obtenir pour chaque variable une tendance globale à travers tous les agents chimiques 2- avoir un portrait des différences entre les agents chimiques. Ainsi, il sera possible d'observer une forte association globale homogène ou hétérogène à travers les substances, une association globale nulle et homogène ou encore une association globale nulle malgré des associations fortes, mais hétérogènes à travers les agents chimiques.

III.5. LOGICIELS

Toutes les analyses ont été effectuées par le logiciel R 3.0.3 statistique (Equipe de développement de base R, Vienne, Autriche). Plusieurs packages *mle4* version 1.1-12 [Bates et al., 2016], *survival* version 2.39-2 [Therneau et al., 2016] et *metafor* version 1.9-8 [Viechtbauer, 2015] ont également été utilisés.

CHAPITRE IV. ANALYSE DESCRIPTIVE ET COMPARAISON DE DEUX BASES DE DONNÉES D'EXPOSITION PROFESSIONNELLE FRANÇAISES COLCHIC ET SCOLA

RÉSUMÉ (FR)

Plusieurs pays ont construit des bases de données d'exposition professionnelle à des produits chimiques. En France, COLCHIC (1987) et SCOLA (2007) coexistent.

Une comparaison descriptive de leur contenu (variables, lieux de travail, agents chimiques et niveaux d'exposition) a été effectuée sur la période 1987 à 2012.

COLCHIC et SCOLA comptent respectivement 841 682 et 152 486 résultats pour 670 et 70 agents chimiques. Elles enregistrent des données portant sur des industries et des métiers similaires et contiennent le même type d'informations. Sur l'ensemble des 17 agents chimiques communs avec plus de 500 résultats chacun, le rapport entre la concentration médiane de COLCHIC et celle de SCOLA est de 3,45 [1,03-14,3] pour la période 2007 – 2012. Cette tendance est confirmée après stratification par industrie, par tâche et par métier, mais est atténuée lorsqu'elle est restreinte à des durées d'échantillonnage similaires.

COLCHIC et SCOLA représentent une source considérable d'informations, mais résultent d'objectifs différents (prévention pour COLCHIC, réglementation pour SCOLA). Les différences potentielles dues aux stratégies devraient être évaluées pour interpréter les données issues de ces deux bases de données.

RÉSUMÉ (EN)

Several countries have built databases of occupational hygiene measurements. In France, COLCHIC and SCOLA co-exist, started in 1987 and 2007 respectively.

A descriptive comparison of the content of the two databases was carried out during the period 1987-2012, including variables, workplaces and agents, as well as exposure levels.

COLCHIC and SCOLA contain respectively 841,682 (670 chemicals) and 152,486 records (70). They cover similar industries and occupations, and contain the same ancillary information. Across 17 common agents with >500 samples, the ratio of the median concentration in COLCHIC to the median concentration in SCOLA was 3.45 [1.03-14.3] during 2007-2012. This pattern remained when stratified by industry, task, and occupation, but was attenuated when restricted to similar sampling duration.

COLCHIC and SCOLA represent a considerable source of information, but result from different purposes (prevention, regulatory). Potential differences due to strategies should be evaluated when interpreting data from these databases.

Keywords: exposure database, risk assessment, sampling strategy, workplace exposure, chemical agents

IV.1. INTRODUCTION

Objective measurements of workers' exposure play an important role in occupational health. First, making a full inventory of chemical agents present in a workplace is a major challenge for risk assessment. This task may be facilitated by the availability of existing measurement data [Bégin et al., 1995]. Moreover, such data enable the implementation of targeted prevention policies by helping prioritize industries and occupations. Lastly, exposure measurements are particularly important as a source of information for carrying out epidemiological studies, for toxicological risk assessment, and to support the development of exposure models.

Aiming at managing large quantities of data generated following the collection of air samples during various occupational health related governmental activities, several large occupational exposure databases have been set up. The oldest contain data dating from the 1970s, such as the Integrated Management Information System (IMIS) from the US Occupational Safety and Health Administration (OSHA) [Stewart et al., 1990] and the MEGA database from the German Institute for Occupational Safety (BIA) [Stamm, 2001]. MEGA and IMIS both include more than 1.6 million data points [Henn et al., 2011; Koppisch et al., 2012]. Others are more recent, such as the NEDB (1986) in the United Kingdom [Burns et al., 1989], SIREP (1996) in Italy [Scarselli et al., 2007], and COLCHIC (1986) [Carton, 1995; Vincent et al., 2001] and SCOLA (2007) in France.

There are many and varied examples of ways in which these data have been used. For instance, Lavoué et al. [2011] estimated levels of exposure to formaldehyde in various industries from the COLCHIC and IMIS data. Lynch et al. [1998] presented an estimation of the number of workers exposed to crystalline silica. Hamm et al. [2011] created a job-exposure matrix for beryllium using IMIS. Pesch et al. [2015] evaluated exposure of welders to hexavalent chromium using the data recorded in MEGA. Scarselli et al. [2012] assessed occupational exposure to hexavalent chromium from the SIREP database in Italy. Coble et al. [2010] created a portrait of chemical agents in the pulp and paper industry using IMIS data.

In France, two occupational exposure data co-exist. COLCHIC was set up in 1986, using data from prevention and research activities; SCOLA was set up in 2006, using data from

compliance activities. COLCHIC and SCOLA come from the same population of workers, but result from different goals, possibly corresponding to different sampling strategies. It is therefore conceivable that they provide different pictures of exposure. No empirical assessment has yet been performed to evaluate this issue.

The objective of this article is to present the first comparative description of the two French databases, COLCHIC and SCOLA, as well as a preliminary assessment of potential differences in exposure levels recorded in these databases.

IV.2. METHODS

IV.2.1. ORIGIN OF COLCHIC AND SCOLA

The first French database of occupational exposure to chemical products (COLCHIC) was established in 1986 at the request of the Commission of Work Accidents and Occupational Diseases. It is managed by the Institut National de Recherche et de Sécurité (INRS). This database contains results of measurements in the air of workplaces by eight French regional health insurance funds, interregional chemical laboratories, and the laboratories of the INRS within the framework of occupational risk prevention. The objectives of setting up COLCHIC were to centralize measurement data from various laboratories, to harmonize workplace sampling and air analysis methods, and to support chemical risk assessment in France. The data in COLCHIC led to several research projects focusing, for example, on exposure to lead, formaldehyde, and mineral fibers [Carton, 1995; Vincent et al., 2002; Kauffer et al., 2007; Lavoué et al., 2006], as well as on exposure prediction models. [Clerc et al., 2015].

The second French occupational exposure database (SCOLA) results from the French requirement that measurements made during assessment of compliance to regulatory occupational exposure limit values (OELV) should be archived in a national register. The measurements stored in SCOLA are performed by certified laboratories, independent of the establishment controlled. The data are used to define priorities for national prevention programs. SCOLA was recently used to update a Web application about *a priori* levels of asbestos dustiness (<http://scolamiante.inrs.fr>).

IV.2.2. CODING OF THE INFORMATION IN COLCHIC AND SCOLA

Ancillary information in COLCHIC can be classified into four major categories: administrative information, sampling conditions, exposure conditions and analytical conditions. Most of the data are coded using standardized classifications specific to COLCHIC, such as task or ventilation, or using national classifications (industries and occupations as coded by the French national systems: classification of activities (NAF)

[INSEE, 2015] and operational directory of occupations and jobs (ROME) [Pôle emploi, 1999], respectively). In 2002, a second version of COLCHIC was developed [Vincent et al., 2001]. A new coding system was established with the modification and/or replacement of existing variables and the creation of new variables in accordance with recommendations of specific work groups [ACGIH-AIHA, 1996; Rajan et al., 1997]. For instance, since 2002, COLCHIC has contained the worker's occupation, and the category of product or process giving rise to the occupational exposure. The original "Individual protection", "Collective protection", "Global ventilation", "Representativeness", and "Type of work place" variables were modified, with either new categories being created or existing ones merged or their definitions altered.

SCOLA was built using the same software system as COLCHIC. The coding of ancillary information is very similar to that in COLCHIC. Establishments cannot refuse to be visited, but the visits are scheduled in advance.

Table IV. summarizes the main differences between the two systems.

IV.2.3. CIRCUMSTANCES LEADING TO RECORDS IN COLCHIC AND SCOLA

The data collected in COLCHIC come from measurement campaigns performed in establishments from the national social security scheme. The choice of targets leading to the measurements in COLCHIC stems from general prevention programs defined by a period of four years by the national health insurance system, as well as from national sampling surveys such as for beryllium (2005-2007), chromium (2012-2014), or styrene (2013-2016). The eight interregional chemical laboratories are assigned to defined geographical areas, and cover companies insured by the national social security scheme (~55% of companies in France; public services and agriculture, small retail and artisans are excluded). Three main situations can trigger a visit to a company. First, a visit can be requested by a safety controller or engineer from the regional health insurance fund. The origin of such a request is highly variable: it may come from an establishment, from an employee representative, or from an occupational physician. Second, visits can be made for research purposes (e.g., industry-wide surveys) in the French prevention network. In this context, companies are usually contacted by the researchers. Third, a visit can occur

within the framework of specific prevention actions performed by each regional laboratory. The laboratory will then identify and contact companies to be visited. Establishments cannot refuse to be visited, but the visits are scheduled in advance.

Table IV.1 Main variables documented in COLCHIC and SCOLA (*Principales variables descriptives documentées dans Colchic et Scola*)

Variable	Category of information ¹	Type of data	Definition
<i>COLCHIC and SCOLA</i>			
Laboratory	E	Category	Organization owning the file
Year	E	Date	Year of reception of intervention request (COLCHIC) or of collecting samples (SCOLA)
File number	E	Category	Number of intervention request file or establishment concerned
SIRET	E	Category	Administrative identification no. of establishment as a geographically localized entity
Establishment – post code – town	E	Category	Address of establishment
NAF	E	Category	Nomenclature code of French Activities (NAF rev. 02, 2008)
Workforce	E	Category	Number of employees in establishment by sector
Collected by	SC	Category	Designation of the entity (COLCHIC) or body (SCOLA) responsible for the sampling campaign in the establishment
Sampling date	SC	Date	Date of sampling in the establishment
Reference / Sample number	SC	Category	Reference of sample given by sampler
Sampling method	SC	Category	Characterization of technique used to take the sample
Sampling device	SC	Category	Characterization of the sampling device (type and dimensions)
Flow rate (L/min)	SC	Continue	Flow rates measured at the start and end of active sampling
Objective**	SC	Category	Type of sample taken (environment or individual; quantitative or qualitative)
Representativeness**	SC	Category	Representativeness of sample indicated by sampler taking into account his or her knowledge of exposure conditions
Type of workplace**	Ex	Category	Brief description of type of workplace in which sample was taken
Global ventilation**	Ex	Category	Brief description of the general ventilation of workplace

Variable	Category of information ¹	Type of data	Definition
Collective protection**	Ex	Category	Description of collective protection equipment used and subjected to sampling
Task	SC	Category	General description of type of task performed during sampling
Occupation*	SC	Category	Identification of type of trade performed by employee (Employment Agency Operational Directory of Trades and Jobs (ROME version 1.2 – 1999))
Personal protective equipment	SC	Category	Description of personal protective equipment used
Duration (min)	SC	Continue	Duration of sampling in minutes
Observations on sampling	SC	Text	Observations recorded during sampling
Laboratory	AC	Category	Laboratory responsible for the analysis requested
Chemical agent	AC	Category	Chemical agent measured during the analysis
Analysis technique	AC	Category	Analysis technique used to quantify chemical agent
<i>COLCHIC alone</i>			
Origin of request	E	Category	Indication of origin of the request for intervention
Reason for request	E	Category	Indication of reason for the request for intervention
Workplace area	E	Text	General localization of place where samples were taken
Estimated production activity*	Ex	Category	Description of production activity of workplace where samples were taken
Ambient temperature	Ex	Category	Range of ambient temperature of workplace when taking the samples
Type of process*	Ex	Category	Description of type of process generating the chemical agents causing exposure
Frequency of exposure*	Ex	Category	Characterization of the frequency of exposure to the category of product or process causing exposure
CAPPROEX*	Ex	Category	Category of product or process causing occupational exposure
Utilization T°*	Ex	Text	Utilization temperature of product causing exposure

E: administrative information on establishment

SC: sampling conditions

Ex: exposure conditions

AC: analytic conditions

* added in 2002 in COLCHIC

** modified in 2002 in COLCHIC

The identification of companies in SCOLA is different. Certified laboratories take measurements at the request of companies in order to fulfill their regulatory obligations. In France, the company director is legally responsible for risk assessment. An initial evaluation is performed by company experts or consultants, which results in a document identifying situations with a potential for overexposure. For each of the situations potentially at risk for which an OELV exists, a measurement strategy has to be elaborated in compliance with the regulation. The current law requires a minimum of three and up to nine samples representative of the workers' exposure in order to verify compliance to an OELV [Le ministre du travail, de la famille, de la solidarité et de la ville, 2009].

IV.2.4. DESCRIPTIVE SUMMARY OF COLCHIC AND SCOLA

The descriptive analysis included data for all agents from 1987 for COLCHIC and from 2007 for SCOLA, up until 2012. Several quality criteria were defined to select data. First, the analysis was restricted to quantitative air measurements performed in the worker's breathing zone (personal sample) or at a fixed point (area sample). Second, units had to be expressed as an air concentration. Measurements with sampling duration less than 60 minutes were identified as "Short term," and all other records as "Long term". Only administrative variables relevant for exposure assessment were included in the dataset. Partial samples (using several sequential samples to evaluate a full day) are treated differently in COLCHIC and SCOLA. In SCOLA, sequential samples are aggregated into a single time weighted average (TWA) value before being recorded. In COLCHIC, sequential samples are entered separately, but information to link them is not well or consistently coded. Using the available information to aggregate partial samples together would require manual processing of each record, which was not possible in the current study.

After data extraction, descriptive statistics were calculated in parallel for both databases. We examined the number of records and their distribution across levels of all variables common to COLCHIC and SCOLA. We also identified the industrial sectors covered by both databases in relation to the French workforce. Finally, we calculated summary statistics of exposure levels for all agents in both COLCHIC and SCOLA with at least 500 records in each database.

IV.2.5. COMPARISON OF EXPOSURE LEVELS BETWEEN COLCHIC AND SCOLA

The comparison of exposure levels between COLCHIC and SCOLA was performed in three parts.

First, for each of the agents having at least 500 measurements in both databases, empirical cumulative distribution curves (ECDF) were created after restriction to personal samples over the common period 2007–2012. Area samples were not included because they can be very different from personal samples, and are infrequent in SCOLA except for asbestos.

Second, to take into account potential confounding due to differences in industries included in the two databases, we identified for each chemical agent a list of industries having at least 20 measurements both in COLCHIC and SCOLA. We chose 20 as a threshold because it corresponded approximately to the number required to cover more than one establishment. We then calculated for each agent and each industry category the proportion of values exceeding the OELV (exceedance), the geometric mean (GM) of recorded concentrations, and the proportion of data lower than the limit of quantification (proportion of non-detects) in COLCHIC and SCOLA. Calculation of the geometric mean when a data was reported as <LOQ was performed using regression on order statistics (ROS) as described in [Helsel, 2012]. In situations where fewer than three detected values were available, GM was estimated by replacing non-detects with $LOQ/2$. For each chemical agent, we evaluated agreement between the two databases across the included industry categories by determining: 1) the mean difference in proportions of data exceeding the OELV, 2) the mean relative difference in GM values and 3) the mean difference in proportions of data reported below LOQ.

In both COLCHIC and SCOLA, the data on the industrial activity are codified by way of a standardized classification using three variables: industry, occupation and task. The previous analysis was also performed using the task and occupation classification systems in place of industry.

Third, we compared exposure levels over the common period 2007–2012 within establishments present in both COLCHIC and SCOLA. For each chemical agent present in

both databases, in at least three common establishments, we calculated, for each of these establishments, the ratio of the geometric mean of measurements in COLCHIC (GMCOLCHIC) to the the geometric mean of measurements in SCOLA (GMSCOLA). Asbestos data were excluded as it was shown that phase-contrast microscopy (used in SCOLA) is not comparable to transmission electron microscopy (used in COLCHIC) [Clerc et al., 2011].

All analyses were performed using the R 3.0.3 statistical software (R Development Core Team, Vienna, Austria).

IV.3. RESULTS

IV.3.1. DESCRIPTIVE SUMMARY OF COLCHIC AND SCOLA

During the period 1987–2012, 841,682 air measurement results from 670 agents were recorded in COLCHIC by eight interregional chemical laboratories, covering 49,145 interventions in 37,569 different companies. From 2007 to 2012, 152,486 air measurement results from 70 agents were recorded in SCOLA by 86 different certified laboratories, covering 37,277 interventions in 7,449 different companies. Eight percent ($n=70,937$) of COLCHIC extracted data were eliminated before analysis according to quality criteria (3.5%, objective was not to characterise the exposure of workers, 4.5% unit was not expressed as an air concentration). All extracted records were used for analysis in SCOLA. Table IV.2 presents the distribution of records in both databases according to categories of the main ancillary variables.

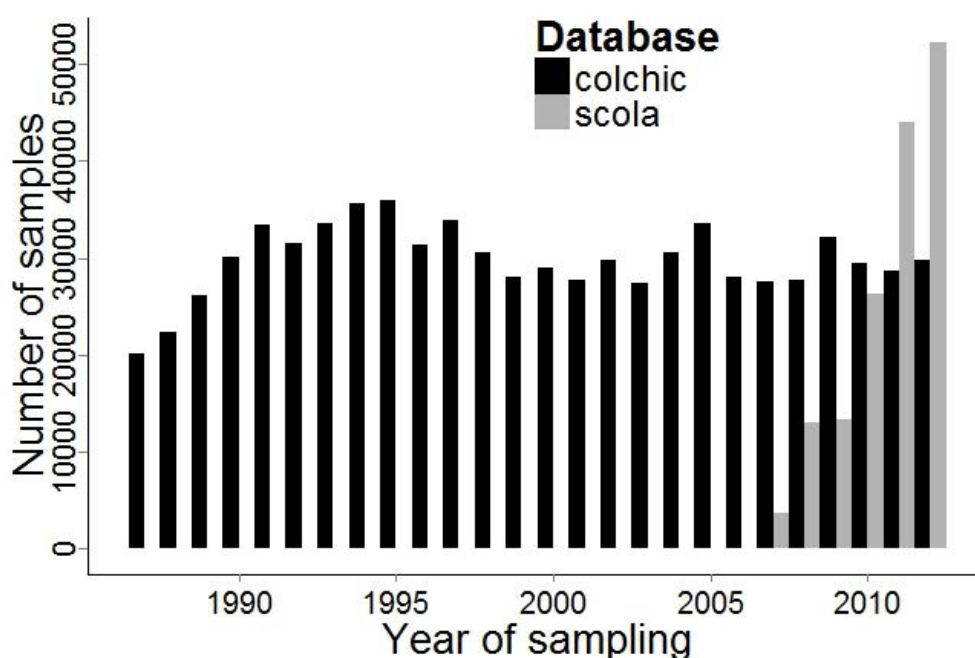


Figure IV.1 Distribution of the number of data per year (Distribution du nombre de mesures par année)

Table IV.2 Distribution of records according to categories of the main variables in COLCHIC and SCOLA (*Distribution du nombre de mesures codées par catégorie de variables communes à COLCHIC et SCOLA*)

Variable of interest	COLCHIC	SCOLA
<i>Personal protection equipment</i>		
Adapted respiratory protection	14% (n = 109,097)	33% (n = 50,021)
No adapted respiratory protection	60% (n = 461,717)	31% (n = 46,928)
Not filled-in	26% (n = 199,931)	36% (n = 55,537)
<i>Collective protection equipment</i>		
Absence of localized capture	47% (n = 360,523)	27% (n = 40,814)
Presence of localized induction (non-enveloping)	28% (n = 216,438)	23% (n = 34,838)
Presence of localized enveloping capture	14% (n = 108,878)	5% (n = 8,005)
Large ventilated booth	1.7% (n = 12,891)	9% (n = 14,203)
Significant distance of employee from emission source	0.2% (n = 1,388)	0.8% (n = 1,338)
System non-operational	0.1% (n = 912)	0.2% (n = 330)
Not filled-in	9% (n = 69,715)	35% (n = 52,958)
<i>Global ventilation</i>		
Presence of general mechanical ventilation	28% (n = 216,893)	45% (n = 68,950)
Absence of mechanical ventilation	35% (n = 268,699)	26% (n = 39,575)
Presence of localized capture and absence of general mechanical ventilation	25% (n = 192,692)	16% (n = 23,963)
System non-operational	0.2% (n = 1,362)	0% (n = 488)
Not filled-in	11.8% (n = 91,099)	13% (n = 19,510)
<i>Type of workplace</i>		
In open air	3.6% (n = 27,679)	9% (n = 12,763)
Confined space	1% (n = 6,947)	35% (n = 53,405)
Booth or control room	1.4% (n = 10,445)	2% (n = 2,972)
Standard industrial premises	73% (n = 563,209)	44% (n = 67,568)
Underground gallery, tunnel	0.5% (n = 3,464)	0% (n = 384)
Other non-classified types	1% (n = 7,057)	7% (n = 10,601)
Archives: partially open	12% (n = 93,926)	0% (n = 0)
Not filled-in	7.5% (n = 58,018)	3% (n = 4,793)
<i>Representativeness</i>		
Development of a sampling or analysis technique	0.3% (n = 1,974)	0% (n = 0)
For comparison with a TWA	30% (n = 231,772)	91% (n = 139,144)
For comparison with a STEL	1.2% (n = 7,719)	9% (n = 13,342)
For determining the exposure for an operation or task	62% (n = 481,505)	0% (n = 0)
For determining the maximum exposure value	0.2% (n = 1,579)	0% (n = 0)
For measuring the concentration at a fixed point without	4% (n = 29,320)	0% (n = 0)

Variable of interest	COLCHIC	SCOLA
concern given for comparison with a limit value		
For evaluating a respiratory protection system	0% (n = 175)	0% (n = 0)
For identifying substances ("screening")	2.3% (n = 16,701)	0% (n = 0)
For determining qualitative exposure and non-representative samples	0.3% (n = 1,974)	0% (n = 0)
<i>Number of employees</i>		
1 to 19 employees	13% (n = 100,985)	14% (n = 21,703)
20 to 49 employees	10% (n = 78,888)	25% (n = 38,035)
50 to 199 employees	14% (n = 106,577)	26% (n = 39,899)
200 or more	8% (n = 64,212)	35% (n = 52,849)
Not documented	55% (n = 420,083)	0% (n = 0)

A median of 31,917 measurements per year were recorded in COLCHIC, relatively stable over the entire study period (Figure IV.1). Measurements appeared in SCOLA in 2007, starting with 3,598 in 2007 and steadily increasing to 52,202 in 2012. While COLCHIC contains a mix of area and personal samples (ratio personal/area 1.44), the compliance related data in SCOLA are much more frequently personal samples (ratio personal/area 2.23, increasing to 34 when excluding data for asbestos).

SCOLA contains a large quantity of asbestos measurements (46% of all data) due to the fact that this contaminant's regulation has warranted measurements for a longer period than other agents. Most samples correspond to monitoring activities for containment and removal of asbestos-containing materials or activities involving interventions on materials or equipment that may release asbestos fibers.

There are also fewer measurements reported as <LOQ in COLCHIC compared to SCOLA (ratio <LOQ/detected 0.31 for COLCHIC vs. 1.13 for SCOLA).

In terms of sampling duration, median sampling duration for personal samples over the entire period covered in COLCHIC was 35 min for short-term samples, and 172 min for long-term samples. For area samples, these numbers were 39 min for short-term samples, and 180 min. for long-term samples. When restricted to the period common to both databases (2007-2012), sampling durations, in COLCHIC and SCOLA, for personal samples were 28 and 15 min respectively for short-term samples, and 215 and 377 min respectively for long term samples. For area samples, these numbers were 41 and 27 min respectively for short term samples, and 195 and 60 min respectively for long-term samples.

A comparison between the distribution of the total number of workers in France per broad industry sector obtained from the National Institute of Statistics and Economic Studies (INSEE) and the distribution of the number of measurements performed in COLCHIC and SCOLA is presented in Figure . The three most frequent industries in COLCHIC are manufacture of metal products, manufacture of rubber and plastic products and manufacture of chemicals and chemical products. They represent 27% of the data. In SCOLA, specialized construction activities, remediation activities and other waste management services and manufacture of other non-metallic mineral products are the three most visited industries. They represent 45% of the data. Globally, 69% of industry categories included in COLCHIC or SCOLA are common to the two databases, representing more than 95% of the data collected.

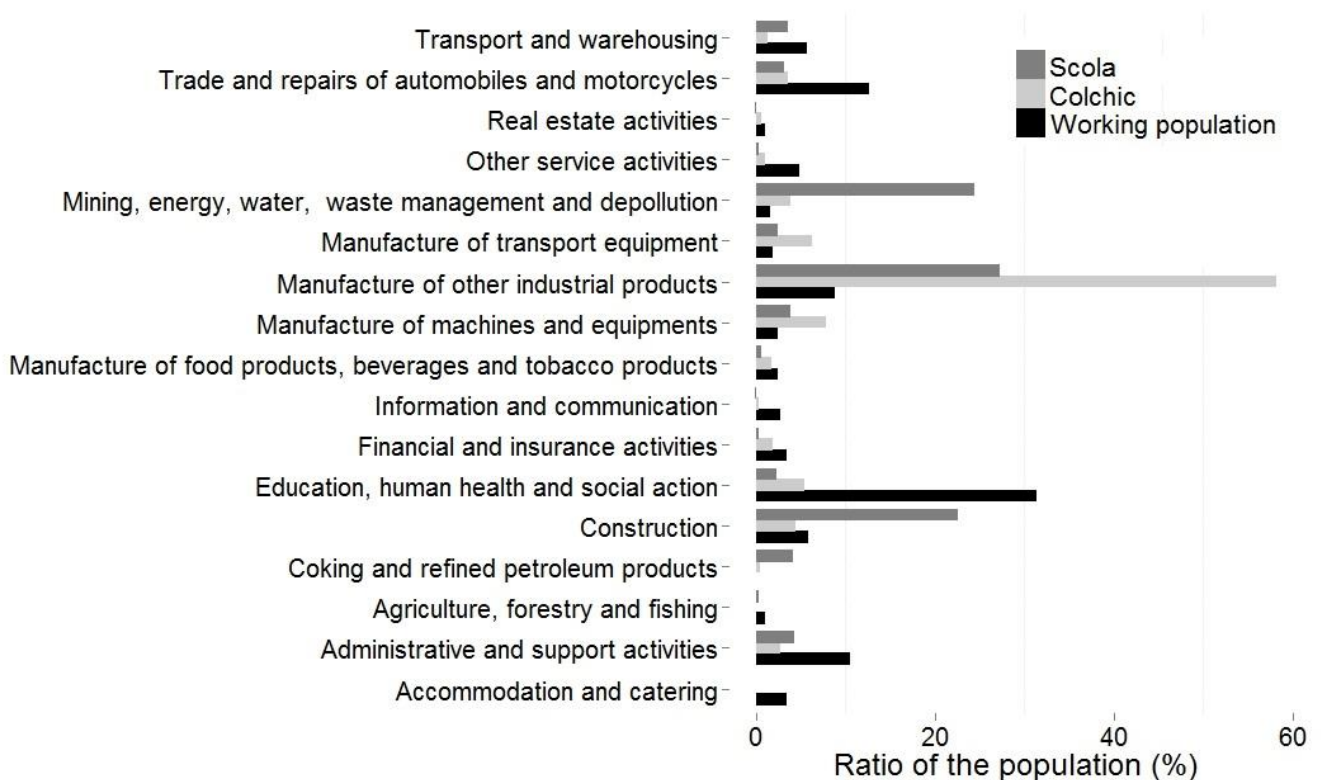


Figure IV.2 Distribution of n for the Industrial sector and comparison with the population by industrial sector in 2012 (*Distribution du nombre de mesures par secteur d'activité et comparaison avec le nombre de travailleurs (en 2012) de ces mêmes secteurs industriels*)

As occupation was recorded only after 2002 in COLCHIC, this variable is “missing” for 62% of records for the total period from 1987 to 2012. Welder, surface treatment worker and

printing press operator are the three most frequent occupations in COLCHIC and represent 8% of the data. In SCOLA, the three main occupations are depolluting treatment officer, production unit supervisor in the chemical and energy industry and machine tool operator. They represent 50% of the measurements. Globally, 70% of occupation categories included in COLCHIC and SCOLA are common to the two databases, representing 46% of records.

The three most frequent tasks coded in COLCHIC, application of air-pressure spraying, machining by removing material and blending operator and controller, represent 13% of the data. In SCOLA, other carcassing work, other special works and other finishing and maintenance works represent 35% of the data. Globally, 65% of task categories included in COLCHIC or SCOLA are common to the two databases, representing 90% of the data collected.

Seventeen chemical agents common to COLCHIC and SCOLA and with >500 measurements represent 38% and 88%, respectively, of records in the databases. The three most measured agents in COLCHIC were inhalable particulate fraction (38,304 personal, 23,476 area), toluene (20,274 personal, 10,935 area) and acetone (19,790 personal, 8,678 area) and in SCOLA, asbestos (45,001 personal, 25,205 area), wood dust (16,055 personal, 485 area) and alveolar particulate fraction, (6,697 personal, 173 area). The proportion of personal measurements (Table) reported below LOQ had a median of 15% across the 17 agents in COLCHIC (ranging from 1%, anhydrous ammonia, to 86%, cristobalite) and of 58% in SCOLA (ranging from 13%, anhydrous ammonia to 92%, cristobalite).

Table IV.3 Descriptive statistics for personal measurements by chemical agent and sample duration in COLCHIC (1987-2012) and SCOLA (2007-2012) (*Statistiques descriptives des mesures d'exposition par agent chimique et par durée de prélèvement pour COLCHIC (1987-2012) et pour SCOLA (2007-2012)*)

Chemical agent	COLCHIC								SCOLA							
	Long term duration				Short term duration				Long term duration				Short term duration			
	Duration		<LOQ ¹ >OELV ²		Duration		<LOQ ¹ >OELV ²		Duration		<LOQ ¹ >OELV ²		Duration		<LOQ ¹ >OELV ²	
	n	(min)	(%)	(%)	n	(min)	(%)	(%)	n	(min)	(%)	(%)	n	(min)	(%)	(%)
		C50 [C25-C75] ³				C50 [C25-C75] ³				C50 [C25-C75] ³				C50 [C25-C75] ³		
1-methoxy- propane-2-ol	2,882	107 [77-207]	23	2	1,251	35 [21-49]	15	5	394	375 [253-431]	45	0.5	282	15 [15-15]	69	2
Acetone	14,093	124 [82-357]	8	0	5,697	37 [25-48]	5	2	740	384 [304-432]	21	0	747	15 [10-15]	37	0
Anhydrous ammonia	218	140 [106-196]	1	18	94	15 [14-39]	3	45	529	415 [370-449]	13	0.8	320	15 [14-15]	52	0
Asbestos	3,588	137 [100-215]	34	0	1,203	18 [10-40]	31	0	35,758	60 [60-60]	70	3	9,243	26 [15-35]	64	16
Benzene	2,104	295 [108-450]	46	10	1,394	32 [20-45]	16	11	3,714	414 [351-445]	74	0.4	334	15 [13-19]	84	1
Butanone	9,041	108 [77-220]	14	2	4,658	37 [26-48]	8	4	828	376 [270-425]	23	0.6	862	15 [10-15]	40	0
Cristobalite	5,900	351 [271-404]	86	3	56	37 [29-44]	82	43	5,606	420 [370-458]	92	0.4	90	17 [9-31]	94	6
Ethylbenzene	5,316	108 [79-233]	24	0.7	3,447	35 [24-47]	10	3	1,954	409 [338-445]	62	0.1	700	15 [7-15]	80	0.4
Methanol	815	138 [86-480]	36	2	454	32 [16-48]	13	12	448	364 [202-420]	60	1	176	15 [5-15]	70	0
N-hexane	2,582	122 [76-360]	28	2	1,808	35 [24-46]	10	4	1,121	418 [372-450]	45	0.7	174	15 [10-15]	70	0

Chemical agent	COLCHIC								SCOLA							
	Long term duration				Short term duration				Long term duration				Short term duration			
	Duration		<LOQ ¹ >OELV ²		Duration		<LOQ ¹ >OELV ²		Duration		<LOQ ¹ >OELV ²		Duration		<LOQ ¹ >OELV ²	
	n	C50 [C25-C75] ³	(%)	(%)	n	C50 [C25-C75] ³	(%)	(%)	n	C50 [C25-C75] ³	(%)	(%)	n	C50 [C25-C75] ³	(%)	(%)
Lead	12,558	175 [120-240]	18	35	1,313	40 [24-50]	15	46	3,456	390 [327-435]	28	15	191	19 [11-30]	68	0
Alveolar particulate fraction	11,962	330 [239-389]	4	5	171	39 [29-50]	16	29	6,587	410 [351-450]	20	3	110	21 [13-39]	72	3
Inhalable particulate fraction	35,366	191 [132-292]	4	0	2,938	40 [26-50]	10	0	3,873	365 [291-413]	19	6	121	25 [15-35]	64	3
Quartz	6,695	349 [269-400]	36	15	65	36 [22-42]	65	32	5,855	421 [372-459]	53	5	91	17 [9-31]	91	7
Toluene	12,955	117 [80-295]	18	10	7,319	35 [23-47]	8	14	2,657	405 [324-444]	46	2	1,424	15 [9-15]	56	5
Wood dust	9,186	282 [194-370]	3	46	159	35 [20-50]	4	83	15,766	380 [325-430]	16	19	289	27 [14-39]	67	12
Xylene	8,530	105 [78-224]	21	1	5,302	35 [23-47]	10	3	2,189	411 [350-446]	50	0.1	984	15 [7-15]	69	0.4

¹Limit of quantification

²Occupational exposure limit value

³Centile 50, 25 and 75

IV.3.2. COMPARISON OF EXPOSURE LEVELS BETWEEN COLCHIC AND SCOLA

The empirical cumulative distribution curves of personal measurements by database and for each agent (Table 3) are presented in Appendix IV.1. Across the 17 chemicals, the ratio of the median concentration in COLCHIC to the median concentration in SCOLA varied between 1.03 (butanone) and 14.3 (n-hexane) with a median of 3.45.

The comparison between personal exposure levels in COLCHIC and SCOLA [2007-2012] in common industries, occupations and tasks is shown in Table IV.44. Results are presented in three parts: 1) the probability of exceeding the OELV, 2) the geometric mean, 3) the proportion of non-detected. Agreement between COLCHIC and SCOLA varied across agents, activity class and metric but, generally indicated higher exposures in COLCHIC compared to SCOLA. Hence, for the industry category, the median of the relative bias of the geometric mean was +38%, the median bias for proportion of non-detects was 3.72 and the median bias for % above the OEL was 0.62%. Because sampling duration was usually higher in SCOLA compared to COLCHIC, and longer samples have often been associated with lower exposure levels, we also performed the comparison restricted to sample durations between 240 and 480 min in both databases, corresponding to between half and 100% of an 8-hour shift (this interval was selected as a compromise between representativeness of full-shift and sufficient sample size). The results are presented in appendix IV.2. They show a decrease, albeit a partial one, in the amplitude of the pattern seen in Table IV.4.

Table IV.4 Comparison of exposure levels in classes of activity common to COLCHIC and SCOLA over the period 2007 to 2012 (*Comparaisons des niveaux d'exposition par classe d'activité commune à COLCHIC et SCOLA sur la période 2007 à 2012*)

Industrial sector

Chemical agent	N ¹	% included SCOLA ²	% included COLCHIC ³	Exceedance			Geometric Mean			Non-detects		
				Bias ⁴	Median SCOLA ⁵	Median COLCHIC ⁶	Bias GM ⁷	Median GM SCOLA ⁸	Median GM COLCHIC ⁹	Bias ND ¹⁰	Median ND SCOLA ¹¹	Median ND COLCHIC ¹²
Acetone	11	65	67	0.37	0	0	80	1.9	12	-0.41	24	14
Toluene	17	48	70	-3.31	0.26	0.94	-1,040	0.32	0.85	7.72	26	30
Butanone	10	67	71	0.35	0	0	-105	3.7	3	9.57	19	28
Xylene	13	49	66	0.68	0	0	43	0.83	2	3.72	21	21
Lead	16	87	87	3.69	5.1	11	38	1.1	3.4	20.9	26	55
Cristobalite	13	89	84	0.62	0	0	53	0.0018	0.0057	-4.50	96	95
Wood (dust)	18	89	97	28.1	19	46	56	0.42	1	-11.3	10	1.1
Quartz	13	89	84	-2.45	4.8	7.4	32	0.005	0.012	-9.57	53	40

Occupation

Chemical agent	N ¹	% included SCOLA ²	% included COLCHIC ³	Exceedance			Geometric Mean			Non-detects		
				Bias ⁴	Median SCOLA ⁵	Median COLCHIC ⁶	Bias GM ⁷	Median GM SCOLA ⁸	Median GM COLCHIC ⁹	Bias ND ¹⁰	Median ND SCOLA ¹¹	Median ND COLCHIC ¹²
Toluene	15	60	47	0.80	0	0	50	0.16	1.4	9.28	26	38
Xylene	10	50	51	1.18	0	0	44	0.34	1	-4.88	48	32
Lead	15	51	66	0.67	19	13	7	3.2	4.9	14.6	19	37
Cristobalite	16	49	71	0.12	1	0.21	-404	0.0017	0.0033	0.03	93	94
Wood (dust)	12	75	94	12.3	24	35	43	0.46	0.79	-9.17	9.2	1.6
Quartz	16	48	71	6.21	4.7	11	39	0.006	0.016	-13.3	47	27

Task

Chemical agent	N ¹	% included SCOLA ²	% included COLCHIC ³	Exceedance			Geometric Mean			Non-detects		
				Bias ⁴	Median SCOLA ⁵	Median COLCHIC ⁶	Bias GM ⁷	Median GM SCOLA ⁸	Median GM COLCHIC ⁹	Bias ND ¹⁰	Median ND SCOLA ¹¹	Median ND COLCHIC ¹²
Toluene	13	57	39	2.36	0	1.7	83	0.13	1.2	-2.28	36	33
Xylene	13	49	47	1.03	0	0	59	0.28	2.4	-8.00	38	26
Lead	17	42	40	-9.14	21	11	-1,25	7.2	8.3	22.68	14	41
Cristobalite	18	32	45	0.69	0	0	-10	0.0010	0.0028	-3.79	95	94
Wood (dust)	10	74	93	15.6	19	37	45	0.39	0.81	-8.77	12	1.1
Quartz	18	32	45	5.21	4.5	6.9	69	0.004	0.013	-18.19	49	34

1: number of classes categories in common with more than 20 measures in the two databases

2: Percentage of SCOLA records included in the comparison (%)

3: Percentage of COLCHIC records included in the comparison (%)

4: Median absolute difference in exceedance fraction of OELV (%)

5: Median of exceedance fraction of OELV across activity categories in SCOLA (%)

6: Median of exceedance fraction of OELV across activity categories in COLCHIC (%)

7: Median relative difference in geometric mean (%)

8: Median of geometric mean values across activity categories in SCOLA (mg/m³ except lead in µg/m³)

9: Median of geometric mean values across activity categories in COLCHIC (mg/m³ except lead in µg/m³)

10: Median absolute difference in proportion of records reported as non detected (%)

11: Median of proportion of records reported as non detected across activity categories in SCOLA (%)

12: Median of proportion of records reported as non detected across activity categories in COLCHIC (%)

The comparison of exposure levels in establishments common to COLCHIC and SCOLA is presented in Table IV.5. Results are presented by agent in three parts: 1) the number of common establishments, 2) the number of measurements in each database, 3) the median ratio of geometric mean across the common establishments. Nine chemical agents met the inclusion criteria. Exposure levels were similar for wood dust and 1-methoxy-2-propyl acetate, showed higher exposure levels in COLCHIC for quartz, cristobalite, lead, toluene, xylene and ethylbenzene [1.5-15], and lower exposure levels in COLCHIC for acetone. Fourteen other chemical agents had one or two establishments in common. The median of their median ratio values was 1.6.

Table IV.5 Comparison of exposure levels among selected chemical agents* in establishments common to COLCHIC and SCOLA over the period 2007 to 2012
(*Comparaison des niveaux d'exposition des établissements communs à COLCHIC et SCOLA pour la période 2007 à 2012*)

Chemical agent	N ¹	COLCHIC		SCOLA		Median ⁴	GM _{COLCHIC} /GM _{SCOLA}	
		n ²	n _{C50} ³	n ²	n _{C50} ³		[C25-C75] ⁵	Range ⁶
Wood (dust)	48	520	9	959	9	1	[0.64 - 1.8]	[0.18 - 20]
Quartz	29	236	7	796	12	1,5	[0.97 - 3.1]	[0.06 - 25]
Cristobalite	28	235	7	775	13	4,1	[1.7 - 6.7]	[0.4 - 27]
Lead	19	234	10	542	12	1,9	[0.32 - 6.2]	[0.05 - 210]
Toluene	11	132	7	137	6	6,5	[0.83 - 23]	[0.012 - 78]
Xylene	6	74	8.5	58	3	3,8	[0.56 - 8]	[0.029 - 3200]
Acetone	5	87	12	38	4	0,38	[0.28 - 3300]	[0.033 - 3900]
Ethylbenzene	4	60	8.5	13	3	15	[9.1 - 26]	[1.6 - 47]
1-Methoxy-2-propyl acetate	3	70	8	20	5	1,1	[0.59 - 2]	[0.053 - 2.9]

* Each chemical agent present in both databases in at least three common establishments (N≥3)

1: Number of establishments common to COLCHIC and SCOLA

2: Number of data points measured in the establishments

3: Median of the number of data points measured in an establishment

4: Median of the ratio of the geometric mean (GM)

5: Centile 25 and 75 of the ratio of the GM

6: Range of the ratio of the GM

IV.4. DISCUSSION

The originality of this study resides in the comparison of the two French databases COLCHIC and SCOLA, with data coming from the same base population. COLCHIC and SCOLA share a common objective of centralising and archiving exposure data to be used as a source of information to support occupational health prevention efforts. Taken together, nearly a million data points are recorded in COLCHIC and SCOLA, putting them among the largest international occupational exposure databases covering an extended time period, such as IMIS in the United States and MEGA in Germany. As such, they represent a remarkable source of information on exposure levels.

IV.4.1. DESCRIPTIVE COMPARISON OF COLCHIC AND SCOLA

COLCHIC and SCOLA include data measured from the same industrial settings and share a similar set of extensive ancillary information. Some COLCHIC variables evolved historically through taking into account recommendations formulated by international work groups [ACGIH-AIHA, 1996; Rajan et al., 1997]. This represents an improvement for future data interpretation and is implemented in SCOLA. However, these changes (fusion, creation of new categories) represent an additional challenge to interpreting historical data mixing the old and new definitions. For instance, for the type of workplace variable, three categories (“Confined space”, “Underground gallery, tunnel” and “Other, non classified types”) were added in 2002 to the four initial categories (“In open air”, “Booth or control room”, “Standard industrial premises” and “Partially open”). When analyzing the relationship between type of workplace and exposure levels, the fact that e.g. “Confined space” samples were coded in another category from 1986 to 2002 (maybe “Standard industrial premises”) might affect the results. It is however hard to predict in advance the amplitude of such phenomenon.

The yearly rate of measurements archived in COLCHIC has been approximately stable at around 30,000 since 1987. SCOLA only started recently and accumulated up to 2012 a less impressive 150,000 data points compared to COLCHIC’s 870,000. However, yearly rates are increasing each year (up to 60,000 in 2012) and should continue to grow as the regulatory requirements to measure exposure and send the results to SCOLA to prove

compliance are progressively enforced in France. SCOLA will therefore increasingly become the major source of information in the next decades.

The number of interventions recorded in COLCHIC and SCOLA by industry is not correlated with the number of employees working in these industries, although most industries have been the object of measurements. This observation is linked to the absence of exposure to chemical products in certain industries and to the fact that measurement campaigns are carried out in priority industries. Regarding industries covered in COLCHIC compared to SCOLA, while the two databases do not share the same distribution of measurements across industries, 95% of measurements are in common categories. Differences in industries and activities covered in the two databases might be related to their different purposes (prevention for COLCHIC and regulation for SCOLA); it renders their joint use valuable since the information collected in one database can offset the absence of information in the other.

Over the period common to both databases, 2007–2012, the duration of sampling was significantly shorter in COLCHIC than in SCOLA. While this difference can be partly due to not being able to aggregate sequential samples in this analysis, it is likely also explained by the two databases' respective objectives. Hence SCOLA is related to regulatory compliance assessment with strict sampling guidelines. In particular, there are very few short term samples in SCOLA reflecting the smaller number of short term OELV in the current French regulations. On the other hand, COLCHIC is the result of measurements taken by insurance hygienists making targeted interventions, more likely focused on specific tasks within a day. In addition, there was a constant increase in sampling duration in COLCHIC over the entire study period, likely reflecting evolving exposure assessment practices and analytical sampling constraints. Since sampling duration has often been associated with exposure levels in large databases of exposure measurements [Lavoué et al., 2011; Clerc et al., 2014; Clerc et al; 2015], with longer sampling time associated with lower exposure levels, these differences should be taken into account when interpreting COLCHIC and SCOLA.

Another distinction between COLCHIC and SCOLA is the types of samples, with personal samples approximately twice more frequent than area samples in COLCHIC. In SCOLA,

area samples are virtually absent when excluding asbestos (which depends on a particularly regulation).

IV.4.2. COMPARISON OF EXPOSURE LEVELS BETWEEN COLCHIC AND SCOLA

While COLCHIC and SCOLA jointly represent a considerable source of exposure information, their data were collected for different purposes. COLCHIC data are measured in companies within industries targeted as potentially problematic, while SCOLA measurements come from companies mandating certified laboratories to measure exposures in their facilities for regulatory compliance assessment. It is therefore questionable whether these differences might lead to differences in exposure levels.

The comparison empirical cumulative distribution curves show that the concentrations recorded in COLCHIC are higher than in SCOLA for the majority of the agents, included in the comparison with a median ratio of the 50th percentile of concentrations around 3. On average, 22% of the measurements of the 17 common chemical agents are lower than the LOQ in COLCHIC versus 51% in SCOLA and 14% are higher than the OELV in COLCHIC versus 5% in SCOLA. Moreover, comparisons of exposure levels within common industry task or occupation categories, also showed a similar, albeit more variable trend, with higher proportions of measurements above the OEL, higher geometric means and lower proportions of measurements reported as below the LOQ. In certain circumstances, exposure levels are very similar in COLCHIC and SCOLA (e.g., acetone, cristobalite) while in others they are very different (e.g., toluene, lead). An analysis restricted to longer sampling time in COLCHIC, limited by a smaller number of agents included, did reduce but did not eliminate the pattern of our original observations. This pattern of higher exposures in COLCHIC was also observed in the comparison of exposure levels within common establishments (the median of the median ratio values $GM_{COLCHIC}/GM_{SCOLA}$ for the nine included agents was 1.9).

Higher exposure levels in COLCHIC are not implausible, since it contains data presumably measured in situations where potential risk was suspected, whereas compliance must be verified in all situations where a contaminant is deemed present in the workplace. Indeed,

a majority (67%) of measurements in COLCHIC were undertaken with the reason for sampling being “possible risk of exposure”. It is therefore possible that COLCHIC would reflect the higher tail of exposure distribution compared to what is found in SCOLA. Very few reports have been published comparing large exposure databases to evaluate whether they provide similar exposure portraits [Vinzents et al., 1994; Clerc et al. 2014; Clerc et al., 2015]. Recently, a descriptive comparison of formaldehyde measurements between MEGA and COLCHIC [Clerc et al. 2015] showed variable differences in median concentrations around a factor from 1.6 to 2 for the five most frequent industries in both databases, interpreted as “comparable exposure levels” by the authors. Similar conclusions were reached by Lavoué et al. [2011] for formaldehyde exposure levels across several industries in the IMIS and COLCHIC databases, with moderately higher levels in COLCHIC (+14%). Peters et al. [2011] assembled and analysed an international database of crystalline silica measurements. In their statistical models, differences between regions corresponded to predicted geometric means within a factor of 4. Our own observations lay within the differences reported in the literature. They warrant caution when jointly interpreting COLCHIC and SCOLA. Multivariate analysis, such as used in Lavoué et al. [2011] or Peters et al. [2011] should allow evaluation of differences between the two databases while taking into account other ancillary information. This would improve the quantification of these differences and help select approaches to better combine them.

IV.4.3. LIMITS OF THE STUDY

Several limits of this study can be noted: firstly, SCOLA is relatively recent. Therefore the comparison was hampered by smaller sample sizes in SCOLA compared to COLCHIC, causing variability in the stratified comparisons. This limitation should be overcome in future updated comparisons, since the number of records entered each year in SCOLA is rapidly growing. Second, the agglomeration of sequential measurements in COLCHIC needs manual intervention, which could not be done for all chemicals agents included in our analysis. This might have affected the comparison of exposure levels between the two databases, although stratified analyses restricted to higher sampling duration in both databases confirmed the main observations, albeit with decreased amplitude.

IV.5. CONCLUSION

COLCHIC and SCOLA are two databases of occupational exposure to chemical products which centralize data stemming from the same population of workers in France. The potential uses of these databases for prevention are numerous: improved information about *a priori* exposure levels when assessing risk within a company, strengthened elaboration of preventive policies, increased retrospective exposure information for epidemiology and worker compensation. The fact that COLCHIC and SCOLA share a similar structure, cover similar agents and may complement each other in terms of workplaces visited, is a strong argument for their joint interpretation. This study also showed differences in exposure levels recorded in COLCHIC and SCOLA, possibly related to the difference in purpose between these two databases (prevention vs compliance). Future comparisons using multivariate analyses [Sauvé et al., 2012], permitting us to simultaneously take into account relevant ancillary variables, should refine this initial portrait and provide insight into how to best interpret the two databases jointly.

CHAPITRE V. ANALYSE PAR MODÉLISATION DE LA BASE DE DONNÉES COLCHIC

V.1. INTRODUCTION

Ce chapitre décrit l'analyse des relations entre les niveaux d'exposition et les variables contextuelles enregistrés dans COLCHIC. Il correspond au premier objectif de la thèse.

Pour cela, nous avons dû définir les périodes à investiguer en nous appuyant sur les descriptions de COLCHIC faites par Vincent [2001] et Mater [2016], présentant les évolutions historiques de la BDEP avec l'apparition en 2002 de 6 nouvelles variables descriptives « Métier », « Activité de production estimée », « Type de procédé », « Fréquence d'exposition », « Catégorie de produit ou de procédé à l'origine de l'exposition professionnelle » et « Température d'utilisation ». A travers ces évolutions et dans le cadre d'exploitations statistiques de la base, deux périodes se dégagent (avant et après 2002) et nécessitent d'être exploitées disjointement, en fonction des descripteurs disponibles. Ainsi, notre analyse par modélisation de COLCHIC s'est portée sur deux périodes 1987–2015 et 2002–2015.

L'une des spécificités de COLCHIC décrite dans le chapitre IV réside dans sa proportion importante de mesures d'ambiance (1/3 environ), les 2/3 restants étant des mesures individuelles. Les études portant sur l'analyse d'agents chimiques de COLCHIC ont systématiquement conclu à des niveaux mesurés en individuel supérieurs aux niveaux mesurés en ambiance de travail [Vincent et Jeandel, 2002 ; Kauffer et Vincent, 2006 ; Lavoué et al., 2006]. Dans ces circonstances, il est apparu important de stratifier les analyses avec d'une part, les données individuelles et d'autre part, les mesures d'ambiance.

L'objectif de ce travail consiste donc à évaluer, pour une gamme d'agents chimiques, l'association entre les niveaux de mesures individuelles et d'ambiance de travail pour les périodes 1987–2015 et 2002–2015 et les variables descriptives associées.

V.2. MÉTHODE

Les éléments présentés dans cette section reposent en grande partie sur les descriptions faites dans le chapitre III. Un résumé des informations majeures complété par une description plus détaillée des spécificités liées à cette analyse par modélisation sont présentés.

V.2.1. PRÉPARATION DES DONNÉES

Pour chacune des périodes 1987 à 2015 et 2002 à 2015, la sélection des données porte sur les mesures d'une durée comprise entre 1 et 720 minutes exprimées en unité d'air et correspondant à des mesures individuelles ou d'ambiance. Seuls les agents chimiques comptant plus de 1 000 résultats pour une période considérée sont conservés. Les résultats renseignés comme « Non détecté » se sont vus attribuer une LoQ empirique.

V.2.2. VARIABLES EXPLOITÉES

Nous avons sélectionné respectivement 10 variables (« Année », « Entreprise », « Durée de prélèvement », « Equipement de protection individuelle », « Tranche d'effectif », « Ventilation générale », « Origine de la demande », « Motif de la demande », « Tâche », « Secteur d'activité ») complétées par 3 variables (« Métier », « Type de procédé », « Fréquence d'exposition ») codées uniquement sur la période 2002-2015 et quatre variables créées « Région », « Type d'échantillonnage », « Durée stratifiée » et « Intervention ».

V.2.3. MODÉLISATION STATISTIQUE POUR CHAQUE AGENT CHIMIQUE

Pour la modélisation, telle que décrite en détail dans le chapitre III, nous avons utilisé des modèles de régression censurés : les modèles Tobit. Ces derniers ont été construits en incluant toutes les variables jugées d'intérêt pour chacune des périodes et en statuant sur la composition de la structure « Aléatoire » puis des structures « Activité économique » et « Tendance temporelle et durée de prélèvement ». Pour la structure « Aléatoire », le choix s'est porté sur la variabilité liée aux interventions. Pour les deux dernières structures, le

choix de la meilleure composition s'est fait à partir des critères de qualité AIC et BIC puis, par période et par type de prélèvement. La qualité du modèle final a été mesurée pour chaque agent chimique avec le coefficient de détermination (R^2).

L'exploitation des estimations de la modélisation s'est faite en calculant un indice d'exposition relatif (RIE) pour les variables de type catégorie. Pour les variables de type continu, les coefficients des modèles correspondent à des pentes et ont directement été exploités.

Deux analyses de sensibilité, permettant de lever certaines interrogations, ont été effectuées lors de cette exploitation. La première analyse porte sur l'impact de la présence d'une structure aléatoire sur les résultats du modèle. En effet, les fonctions qui permettent d'intégrer une structure aléatoire dans un modèle Tobit sont encore à la pointe de la recherche et ne sont pas encore complètement standardisées. Pour évaluer cet impact, nous avons vérifié que nous obtenions des résultats similaires avec un modèle plus classique sans effet aléatoire. La seconde analyse a permis d'étudier l'impact sur les associations lorsque pour un agent chimique on a très peu de données de courte durée. En effet, pour plusieurs combinaisons proposées pour la structure « Temporelle temporelle et durée de prélèvement », la variable « Type d'échantillon » est en interaction avec la durée de prélèvement et/ou l'année. Afin de nous assurer qu'en retirant les données « Courte durée », lorsqu'il y en avait peu (<20 %), nous n'allions pas introduire d'artefacts, nous avons contrôlé que les résultats, en ne conservant que des mesures « Longue durée », étaient similaires.

V.2.4. MÉTA ANALYSE

A l'issue de la modélisation, les coefficients estimés ont été synthétisés par la Méta analyse. Un coefficient global, assimilé à une pondération des coefficients de tous les agents chimiques, est calculé. Il est ensuite exprimé sous la forme de RIE pour les variables de type catégorie. Pour les variables continues, le coefficient issu de la Méta analyse exprime la pente moyenne à tous les agents chimiques. En complément, après avoir identifié les familles regroupant un nombre significatif d'agents chimiques, une Méta analyse stratifiée par famille d'agents (solvants et métaux) a été effectuée.

La représentation graphique sous forme de Forest plots illustre la variabilité interagents chimiques.

V.3. RÉSULTATS

V.3.1. ANALYSE DESCRIPTIVE

Sur la période 1987-2015, l'extraction COLCHIC comptabilise 929 700 données. Cela correspond à 35 952 interventions concernant 18 777 établissements différents. Sur la période 2002–2015, 441 567 mesures correspondent à 15 930 interventions dans 10 049 établissements. Respectivement 8 % (n = 83 932) et 16 % (n = 72 366) des données extraites pour chacune des 2 périodes (1987-2015 et 2002-2015) sont éliminées avant l'analyse au regard des critères définis. COLCHIC compte 59 % de mesures individuelles contre 41 % de mesures d'ambiance (Figure V.1).

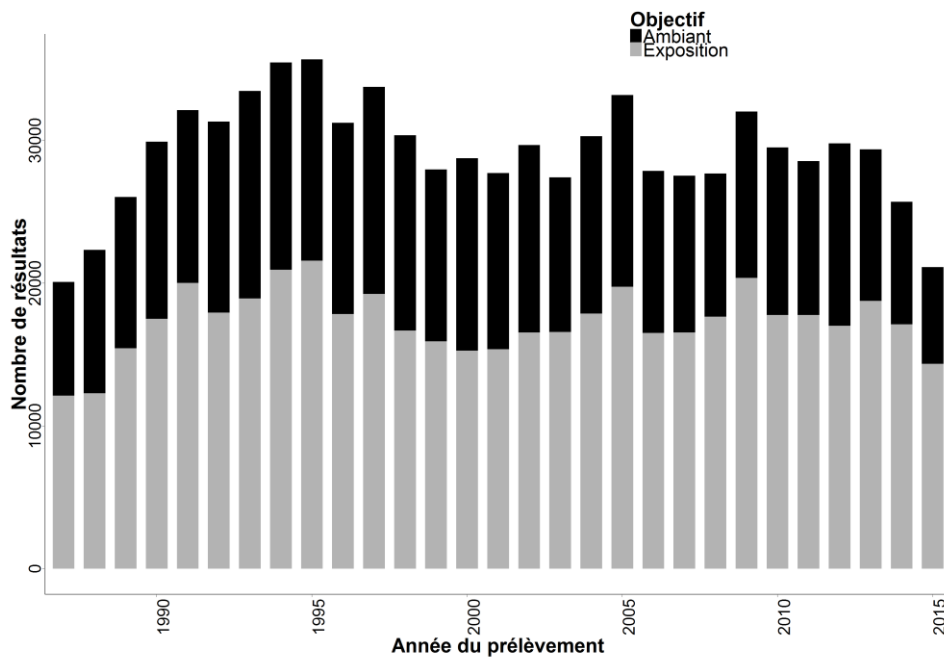


Figure V.1 Distribution annuelle du nombre de mesures individuelles et d'ambiance dans COLCHIC

Tableau V.1 Distribution du nombre de données (%) par catégories de variables incluses dans les modèles pour les périodes 1987-2015 et 2002-2015 par mesure individuelle et par mesure d'ambiance de travail

Variable	Description / Type	1987-2015		2002-2015	
		Individuel	Ambiant	Individuel	Ambiant
Effets fixes					
Région	Région où les prélèvements ont été effectués / Nominal (13 catégories)				
	Poitou Charentes Limousin Aquitaine (PCLA)	38303 (8,8)	22060 (8,1)	15311 (8,3)	7421 (7,8)
	Auvergne Rhône Alpes (ARA)	51144 (12)	28409 (10)	20827 (11)	10487 (11)
	Bourgogne-Franche-Comté (Bourgo_comte)	34721 (8)	22171 (8,1)	14969 (8,1)	8007 (8,4)
	Bretagne	16405 (3,8)	11501 (4,2)	6967 (3,8)	3722 (3,9)
	Champagne Ardenne Lorraine Alsace (CALA)	31774 (7,3)	22440 (8,2)	13755 (7,5)	7744 (8,1)
	Centre	32489 (7,4)	20442 (7,5)	15066 (8,2)	7236 (7,6)
	Département d'Outre-Mer et Corse (DOM Corse)	24143 (5,5)	16227 (6)	10167 (5,5)	6350 (6,6)
	Ile de France (IDF)	54542 (13)	34052 (13)	23418 (13)	11598 (12)
	Midi-Pyrénées Languedoc Roussillon (MPLR)	69874 (16)	42314 (16)	29155 (16)	15409 (16)
	Nord Pas-de-Calais Picardie (Nord)	18022 (4,1)	12138 (4,5)	7610 (4,1)	4488 (4,7)
	Normandie	23637 (5,4)	15806 (5,8)	10651 (5,8)	5397 (5,6)
	Pays de la Loire (P_Loire)	25566 (5,9)	15316 (5,6)	10809 (5,9)	5022 (5,2)
	Provence Alpes Côte d'Azur (PACA)	15509 (3,6)	9197 (3,4)	5825 (3,2)	2792 (2,9)
Motif de la demande	Information sur le motif de la demande d'intervention / Dichotomique (2 catégories)				
	Risque d'exposition	301362 (69)	180909 (66)	149659 (81)	75764 (79)
	Autre raison	134767 (31)	91164 (34)	34871 (19)	19909 (21)
Origine de la demande	Information sur l'origine de la demande d'intervention / Nominal (4 catégories)				
	Organisme prévention	256836 (59)	160573 (59)	105618 (57)	58912 (62)
	Entreprise	83989 (19)	59141 (22)	39449 (21)	22180 (23)

Variable	Description / Type	1987-2015		2002-2015	
		Individuel	Ambiant	Individuel	Ambiant
	Corps médical	78823 (18)	39318 (14)	35357 (19)	12818 (13)
	Autre origine	16481 (3,8)	13041 (4,8)	4106 (2,2)	1763 (1,8)
Ventilation globale	Description de la ventilation générale sur le lieu de travail / Nominal (5 catégories)				
	Non mais présence d'un captage localisé	130904 (30)	72963 (27)	60223 (33)	28647 (30)
	Oui et présence d'un captage localisé	60855 (14)	39669 (15)	29320 (16)	15093 (16)
	Oui	51524 (12)	36976 (14)	23682 (13)	14685 (15)
	Non	136063 (31)	86669 (32)	57508 (31)	29422 (31)
	Information non relevée	56783 (13)	35796 (13)	13797 (7,5)	7826 (8,2)
Tranche d'effectif	Nombre de travailleurs dans l'établissement visité / Nominal (5 catégories)				
	1-19 salariés	60359 (14)	42396 (16)	39221 (21)	23725 (25)
	20-49 salariés	52998 (12)	29386 (11)	43266 (23)	20230 (21)
	50-199 salariés	69963 (16)	39201 (14)	58622 (32)	28615 (30)
	200 salariés ou plus	33422 (7,7)	21811 (8)	27220 (15)	15061 (16)
	Information non relevée	219387 (50)	139279 (51)	16201 (8,8)	8042 (8,4)
Equipement de protection individuelle	Description de l'appareil de protection respiratoire du travailleur / Nominal (3 catégories)				
	Non adapté	267450 (61)	160290 (59)	128952 (70)	58239 (61)
	Adapté	79638 (18)	31220 (11)	44612 (24)	13035 (14)
	Information non relevée	89041 (20)	80563 (30)	10966 (5,9)	24399 (26)

Variable	Description / Type	1987-2015		2002-2015	
		Individuel	Ambiant	Individuel	Ambiant
Type de procédé*	Description du type de procédé à l'origine de l'exposition / Nominal (4 catégories)				
	Ouvert	-	-	146337 (79)	67468 (71)
	Semi-ouvert	-	-	20738 (11)	14565 (15)
	Fermé	-	-	13944 (7,6)	10562 (11)
	Autre type de procédé	-	-	3511 (1,9)	3078 (3,2)
Fréquence d'exposition*	Fréquence d'exposition du travailleur / Nominal (4 catégories)				
	Permanent	-	-	104522 (57)	52891 (55)
	Fréquent	-	-	50880 (28)	21785 (23)
	Occasionnel	-	-	25549 (14)	15743 (16)
	Autre fréquence d'exposition	-	-	3579 (1,9)	5254 (5,5)
Durée d'échantillonnage	Durée du prélèvement / Dichotomique (2 catégories)				
	Longue durée	339943 (78)	223867 (82)	153638 (83)	85900 (90)
	Courte durée	96186 (22)	48206 (18)	30892 (17)	9773 (10)
Année	Année du prélèvement / Continue				
Effets aléatoires					
Intervention	Triptyque Année / laboratoire / Numéro de Dossier COLCHIC				

*Variable ajoutée en 2002

Le Tableau V.1 présente la distribution du nombre de mesures enregistrées dans COLCHIC selon les catégories de variables retenues dans l'analyse pour les agents chimiques comptant plus de 1 000 résultats, par période et par type de prélèvement.

Les Annexes V.1 et V.2 définissent par agent chimique sa famille d'appartenance, le pourcentage de données inférieures à la LoQ, la répartition des données « Courte durée » et « Longue durée » par période et par type d'échantillonnage. 86 et 50 agents chimiques, respectivement pour les périodes 1987-2015 et 2002-2015 comptant plus de 1 000 mesures individuelles et/ou d'ambiance, sont recensés. Ils représentent 84 % et 76 % des données conservées (17 et 13 métaux et leurs composés ; 37 et 23 solvants organiques ; 32 et 14 autres agents chimiques). Quelle que soit la période considérée, les 4 agents chimiques les plus fréquemment mesurés sont les poussières inhalables ($n_{87-15} = 66\,838$; $n_{02-15} = 29\,481$), le toluène ($n_{87-15} = 33\,136$; $n_{02-15} = 12\,734$), l'acétone ($n_{87-15} = 32\,836$; $n_{02-15} = 14\,321$) et le fer ($n_{87-15} = 24\,201$; $n_{02-15} = 11\,817$). La proportion médiane, à travers les agents de résultats inférieurs à la LoQ, est respectivement pour les mesures individuelles et d'ambiance de 22,5 % [min : 2,7 % - max : 92 %] et 26,5 % [4,8 % - 94 %] sur 1987-2015 et de 28 % [1,8 % - 93 %] et 26,5 % [4,7 % - 93 %] sur 2002-2015.

V.3.2. MODÉLISATION DE COLCHIC

Concernant la multicolinéarité, tous les couples de variables indépendantes ont une corrélation inférieure au seuil 0,7.

Après avoir défini les variables à tester pour le modèle final de chaque période, différentes combinaisons pour les structures « Activité économique » et « Tendances temporelles et durée de prélèvement » ont été testées. Globalement, pour tous les agents chimiques et quelle que soit la période, un message cohérent entre les deux critères de qualité du modèle est apparu. Les Tableaux V.2 et V.3 présentent, respectivement pour les structures « Activité économique » et « Tendances temporelles et durée de prélèvement », les différentes combinaisons modélisées par période pour tous les agents chimiques. Les résultats exprimés pour les mesures individuelles et pour les mesures d'ambiance correspondent à la proportion entre le nombre de fois où l'AIC et le BIC sont les plus bas pour un agent chimique et le nombre total d'agents chimiques. Ils

permettent ainsi d'identifier la structure qui sera retenue pour le modèle final. Dans le cas de la structure « Activité économique », on notera que le modèle avec le secteur d'activité présente majoritairement les meilleurs BIC (moins pertinent pour AIC) quelle que soit la période. Concernant la structure « Tendances temporelles et durée de prélèvement », le constat d'une structure prépondérante est moins marqué. Même si la structure (8) montre les résultats les plus satisfaisants, elle ne représente pas la majorité des situations. Dans ce contexte, il est décidé de retenir pour le modèle final la structure la plus complète (9) puisqu'elle intègre toutes les interactions possibles.

Tableau V.2 Choix de la structure « Activité économique »

Combinaison de structure « Activité économique »	Période 1987-2015				Période 2002-2015			
	AIC		BIC		AIC		BIC	
	Ind ⁽¹⁾	Amb ⁽¹⁾	Ind	Amb	Ind	Amb	Ind	Amb
SECTEUR ACTIVITE	47	62	94	94	35	46	96	86
TACHE	53	38	6	6	22	8	2	3
METIER	-	-	-	-	43	46	2	11

⁽¹⁾ Proportion (%) du nombre de fois où l'AIC ou le BIC est le plus bas pour un agent chimique par rapport au nombre total d'agents chimiques évalués

Tableau V.3 Choix de la structure « Tendances temporelles et durée de prélèvement »

Combinaison de structure « Tendances temporelles et durée de prélèvement » ⁽¹⁾	Période 1987-2015				Période 2002-2015			
	AIC		BIC		AIC		BIC	
	Ind ⁽²⁾	Amb ⁽²⁾	Ind	Amb	Ind	Amb	Ind	Amb
(1)	3	6	20	14	5	6	16	14
(2)	3	3	2	3	7	6	9	0
(3)	18	18	5	8	16	19	9	3
(4)	2	3	5	8	0	3	7	6
(5)	0	3	2	5	0	0	0	0
(6)	9	3	20	30	5	17	14	43
(7)	8	11	8	5	9	3	7	6
(8)	35	24	33	21	34	25	39	26
(9)	23	29	8	8	25	22	0	3

⁽¹⁾ Structures « Tendances temporelles et durée de prélèvement » présentées au paragraphe 6.3 du chapitre « Méthodes »

⁽²⁾ Proportion (%) du nombre de fois où l'AIC ou le BIC est le plus bas pour un agent chimique par rapport au nombre total d'agents chimiques évalués

A travers tous les agents chimiques exploités, les proportions médianes de la variabilité expliquée par le modèle complet, par période 1987-2015 / 2002-2015 et par type de prélèvement Individuel / Ambiant, sont présentées dans le Tableau V.4.

Tableau V.4 Coefficient de détermination par période et par type de prélèvement

Période	Type de prélèvement	Coefficient de détermination (R ²)		
		Min (AC)	Médiane	Max (AC)
1987-2015	Individuel	14 % (Poussières de bois)	33 %	69 % (Benzoeacephenanthrylene)
	Ambiant	18 % (Equivalent_NCl ₃)	34 %	55 % (Ammoniac)
2002-2015	Individuel	14 % (Poussières de bois)	33 %	59 % (Tétrachloroéthylène)
	Ambiant	20 % (Poussières de bois)	33 %	49 % (Alcool isopropylique)

(AC) Agent chimique correspondant au coefficient de détermination

Le Tableau V.5 montre, pour les mesures individuelles et les mesures d'ambiance, l'association de tous les effets fixes avec les niveaux stratifiés par période pour tous les agents chimiques. Les Annexes V.3 et V.4 présentent leur association stratifiée par période et par famille d'agents chimiques. Les prélèvements de longue durée ont des niveaux globalement deux fois plus faibles que les prélèvements de courte durée quels que soient le type de prélèvement, la famille de l'agent chimique et l'année [0,46 – 0,64]. Annuellement, les concentrations diminuent. Cette décroissance est plus marquée : 1- sur la période la plus récente (8 % et 10 % respectivement pour les mesures de courte et longue durée) comparée à la période totale (4 % et 5 %) quel que soit le type de prélèvement 2- pour les solvants (10 %) comparés aux métaux (7 %) de 2002 à 2015. D'une région à une autre, les niveaux enregistrés sont hétérogènes. Lorsque le motif de la demande est lié à un risque d'exposition, les niveaux en individuel sont plus importants (2 % pour 1987-2015 et 14 % pour 2002-2015). Cette tendance n'est pas vérifiée pour les mesures d'ambiance ni pour les métaux 1987-2015. Lorsqu'un organisme de prévention est à l'origine de la demande, les niveaux mesurés en individuel sont globalement plus élevés. Cette tendance n'est toutefois pas constatée pour les familles ni pour les mesures d'ambiance. La présence d'une ventilation générale combinée ou non à un captage localisé coïncide avec des niveaux généralement plus élevés, excepté pour les métaux. Les niveaux les plus faibles sont enregistrés en individuel dans les établissements de plus petite taille (1 à 19 salariés) et en ambiant dans les entreprises de plus de 200 salariés. Un

travailleur équipé de protection individuelle adaptée est plus exposé que lorsque l'équipement n'est pas adapté [39 % - 86 %]: cette tendance est accentuée pour les métaux [87 % - 125 %]. Un procédé ouvert engendre systématiquement un niveau plus important. Une exposition permanente ou fréquente est globalement liée à des concentrations plus fortes que lorsqu'elle est occasionnelle [4 % - 12 %].

Les effets résumés Tableau V.5 représentent les Méta tendances et ne permettent pas d'évaluer l'homogénéité à travers les agents chimiques. Un Forest plot montre, à travers le RIE, la comparaison entre une catégorie et celle définie comme référence pour chaque agent chimique (Figure V.2). Les carrés noirs représentent le RIE estimé par substance. Sa taille est proportionnelle à son poids statistique. Les lignes horizontales correspondent à l'intervalle de confiance (IC) à 95 %. Enfin, le losange noir représente l'estimation globale du RIE issue de la Méta analyse et son IC 95 %. Toutes les représentations graphiques de notre étude, stratifiées par type de prélèvement, par période et par famille, sont disponibles à l'adresse <http://expostats.ca/gautier/annexe1.html>. Plusieurs tendances, à travers les agents chimiques étudiés, peuvent être observées : 1- tendance générale supérieure à la référence : par exemple, en individuel sur 2002-2015, les RIE de la catégorie utilisation d'équipement de protection individuelle adapté sont supérieurs à la référence (non adapté) pour 39/40 agents chimiques [1,03-5,86] (Figure V.2, A) 2- tendance générale inférieure à la référence : par exemple, pour 33/40 agents, en individuel sur 2002-2015, les niveaux sont plus faibles [0,27-0,99] lors de la mise en œuvre d'un procédé fermé comparé à un procédé ouvert (Figure V.2, B) 3- tendance générale proche de la référence : par exemple, en individuel sur 2002-2015, lors d'une exposition fréquente à un produit chimique comparé à une exposition permanente, les niveaux sont relativement homogènes à travers les agents : 18/40 agents chimiques ont leur RIE compris entre [0,90-1,10] (Figure V.2, C).

Les analyses complémentaires réalisées, 1- avec et sans structure aléatoire 2- avec les données de courte durée si supérieures à 20 % et sans les données de courte durée, ne montrent pas, de manière générale, de variations notables des Méta coefficients (Annexes V.5 à V.8).

Tableau V.5 Synthèse des indices d'exposition relative (RIE) et de leur intervalle de confiance à 95 % (IC 95 %) par Méta stratifiée par période 1987-2015 et 2002 – 2015 et par type de prélèvement

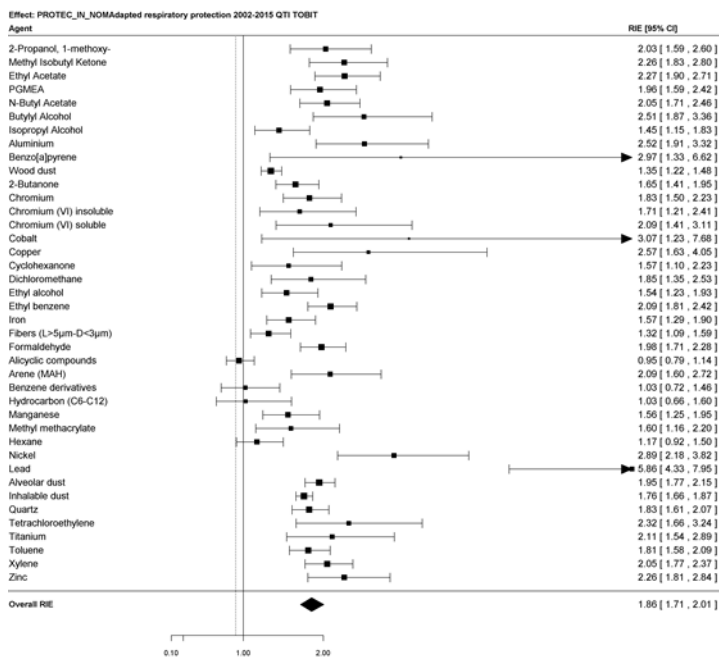
Variable Catégorie	Individuel		Ambiant	
	1987–2015 N=63 RIE (IC 95 %)	2002–2015 N=40 RIE (IC 95 %)	1987–2015 N=59 RIE (IC 95 %)	2002–2015 N=31 RIE (IC 95 %)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	35 % [30-39]	37 % [31-42]	30 % [17-41]	40 % [-72-79]
Longue durée (≥ 60 minutes)	4 % [3-5]	4 % [3-5]	5 % [4-6]	5 % [4-6]
Ratio Conc(240)/Conc(30) en 1987 / 2002^(1,2)	0,62 (0,54;0,72)	0,56 (0,51;0,63)	0,46 (0,33;0,65)	0,64 (0,43;0,95)
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,54 (0,49;0,59)	0,54 (0,50;0,58)	0,53 (0,46;0,63)	0,49 (0,26;0,95)
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	5 % [4-6]	10 % [8-11]	5 % [3-8]	10 % [-3-21]
Longue durée (≥ 60 minutes)	5 % [4-6]	8 % [6-10]	4 % [3-5]	8 % [6-9]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,04 (0,95;1,13)	1,14 (1,01;1,29)	1,07 (0,97;1,19)	1,00 (0,90;1,12)
Bourgo_comte	0,97 (0,89;1,05)	1,00 (0,91;1,09)	1,00 (0,92;1,08)	0,98 (0,87;1,09)
Bretagne	0,94 (0,83;1,07)	0,86 (0,71;1,03)	0,92 (0,79;1,07)	0,99 (0,81;1,20)
CALA	1,00 (0,89;1,11)	1,14 (1,01;1,27)	1,08 (0,96;1,21)	1,11 (0,97;1,27)
Centre	1,02 (0,92;1,12)	1,08 (0,95;1,24)	1,06 (0,95;1,18)	1,11 (0,94;1,30)
DOM Corse	0,87 (0,78;0,98)	0,98 (0,87;1,11)	1,01 (0,87;1,16)	1,03 (0,86;1,24)
IDF	1,00 (0,92;1,08)	1,03 (0,93;1,14)	0,95 (0,86;1,06)	0,93 (0,84;1,03)
MPLR	0,99 (0,90;1,07)	1,04 (0,95;1,14)	1,03 (0,92;1,15)	1,09 (0,97;1,24)
Nord	1,06 (0,95;1,18)	1,29 (1,07;1,54)	1,10 (0,93;1,30)	1,09 (0,85;1,40)
Normandie	1,15 (1,04;1,27)	1,10 (0,95;1,26)	1,15 (1,04;1,27)	0,97 (0,85;1,11)
P_Loire	0,90 (0,80;1,00)	0,90 (0,78;1,05)	0,91 (0,79;1,04)	0,92 (0,78;1,08)
PACA	1,25 (1,08;1,44)	1,05 (0,87;1,26)	1,13 (1,02;1,25)	1,03 (0,85;1,25)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Autre raison	0,98 (0,93;1,03)	0,86 (0,78;0,95)	1,08 (1,01;1,17)	0,97 (0,87;1,09)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Entreprise	0,91 (0,87;0,97)	0,89 (0,82;0,98)	0,92 (0,86;0,98)	0,97 (0,89;1,05)
Corps médical	1,02 (0,94;1,11)	0,99 (0,87;1,12)	0,92 (0,85;1,00)	0,99 (0,88;1,12)
Autre origine	0,88 (0,71;1,11)	0,81 (0,63;1,04)	1,13 (0,92;1,37)	1,18 (0,70;2,02)

Variable Catégorie	Individuel		Ambiant	
	1987-2015 N=63 RIE (IC 95 %)	2002-2015 N=40 RIE (IC 95 %)	1987-2015 N=59 RIE (IC 95 %)	2002-2015 N=31 RIE (IC 95 %)
Ventilation générale				
Non mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,07 (1,00;1,13)	0,98 (0,92;1,05)	1,20 (1,08;1,33)	1,09 (0,98;1,22)
Oui	1,14 (1,08;1,21)	1,08 (1,00;1,17)	1,18 (1,06;1,32)	1,08 (0,94;1,24)
Non	1,06 (1,01;1,12)	0,97 (0,90;1,04)	0,98 (0,93;1,04)	0,84 (0,76;0,94)
Information non relevée	0,99 (0,91;1,07)	0,88 (0,80;0,98)	1,00 (0,91;1,10)	0,89 (0,78;1,02)
Tranche d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,16 (1,07;1,26)	1,10 (1,00;1,22)	1,00 (0,89;1,11)	1,11 (0,99;1,23)
50-199 salariés	1,07 (0,99;1,17)	1,04 (0,94;1,16)	0,97 (0,89;1,05)	0,96 (0,87;1,07)
200 salariés ou plus	1,14 (1,02;1,27)	1,05 (0,91;1,21)	0,98 (0,88;1,09)	0,93 (0,84;1,02)
Information non relevée	1,10 (1,02;1,20)	1,05 (0,95;1,17)	1,03 (0,94;1,13)	1,02 (0,87;1,18)
Equipelement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,67 (1,57;1,77)	1,86 (1,71;2,01)	1,62 (1,46;1,81)	1,39 (1,22;1,59)
Information non relevée	0,99 (0,90;1,09)	0,96 (0,86;1,08)	1,04 (0,96;1,12)	0,89 (0,80;1,00)
Type de procédé				
Ouvert	-	1,00 (Réf.)	-	1,00 (Réf.)
Semi-ouvert	-	0,72 (0,66;0,79)	-	0,96 (0,88;1,06)
Fermé	-	0,73 (0,66;0,81)	-	0,92 (0,81;1,04)
Autre type de procédé	-	0,80 (0,65;0,99)	-	0,89 (0,73;1,09)
Fréquence d'exposition				
Permanent	-	1,00 (Réf.)	-	1,00 (Réf.)
Fréquent	-	1,06 (1,00;1,12)	-	1,06 (0,99;1,13)
Occasionnel	-	0,88 (0,82;0,95)	-	0,96 (0,86;1,08)
Autre fréquence d'exposition	-	0,89 (0,74;1,07)	-	0,78 (0,66;0,93)

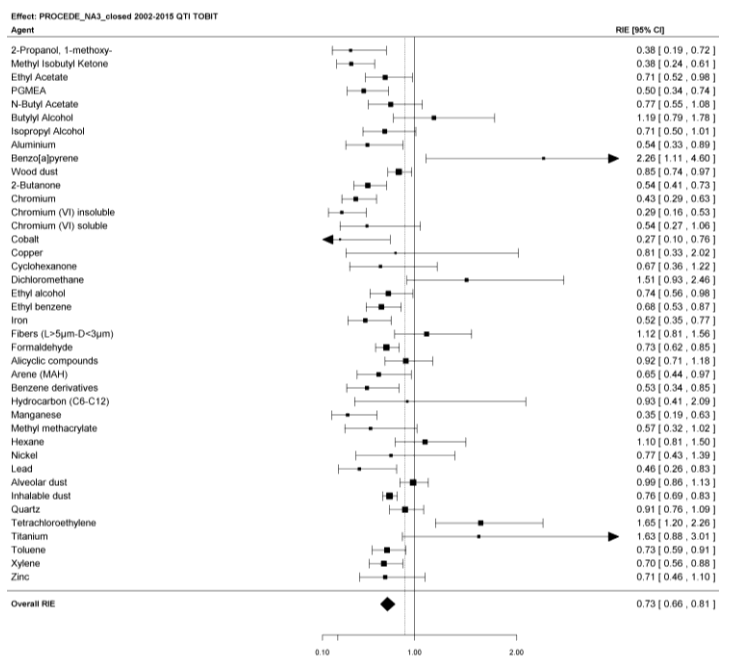
⁽¹⁾ Données calculées avec 39 et 23 agents chimiques respectivement sur 1987 – 2015 et sur 2002 – 2015 pour les individuels et 17 et 2 agents chimiques pour les ambiants

⁽²⁾ Lire 1987 pour la période 1987-2015 et 2002 pour la période 2002-2015

(A)



(B)



(C)

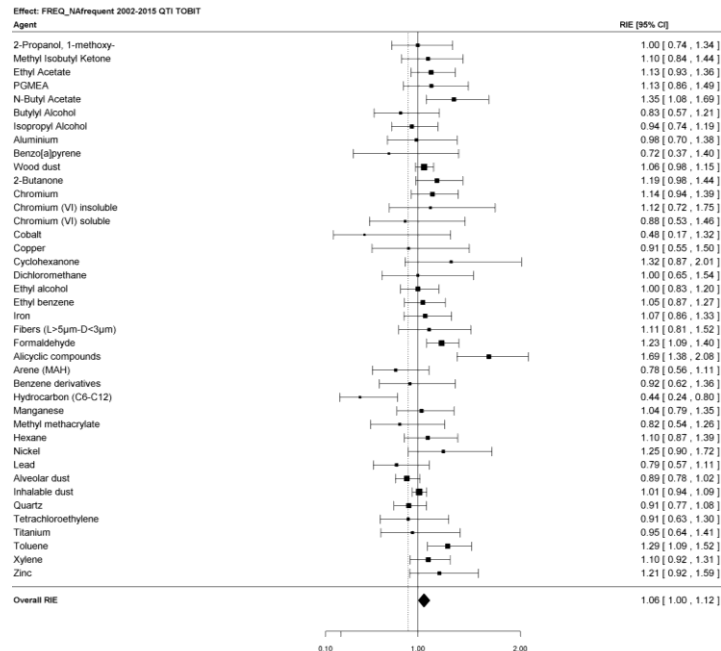


Figure v.2 : RIE issus de la Méta analyse par agent chimique mesuré en individuel sur la période 2002-2015 pour les catégories : A- Equipement de protection individuelle adapté (Ref : non adapté) B- Type de procédé fermé (Ref : ouvert) C- Exposition fréquente (Ref : permanent)

V.4. DISCUSSION

V.4.1. MODÉLISATION STATISTIQUE

A notre connaissance, cette étude représente le premier effort pour examiner globalement l'association entre les niveaux de concentration enregistrés dans COLCHIC et les variables décrivant les circonstances de prélèvement. Les analyses portent sur 68 % des données de COLCHIC et ont été stratifiées 1- par période 1987–2015 et 2002–2015 afin de tenir compte des évolutions historiques de COLCHIC 2- par type de prélèvement « individuel » et « ambiant » 3- pour tous les agents chimiques et par famille de solvants et de métaux, la Méta analyse a permis de présenter une image générale des effets de chaque déterminant d'exposition (Tableau V.5) et, leur homogénéité à travers tous les agents chimiques est présentée sous forme de Forest plots sur la page : <http://expostats.ca/gautier/annexe1.html>.

Les pourcentages de la variabilité totale expliqués par les effets fixes pour chaque strate [14 % - 69 %] sont comparables avec ceux de la littérature pour des études similaires [Burstyn et Teschke, 1999 ; Lavoué et al., 2005 ; Sarazin et al., 2016], respectivement [20 % - 70 %], [35 % - 57 %] et [16 % - 24 %] et confortent nos analyses.

La durée du prélèvement est un prédicteur majeur des niveaux individuels de COLCHIC. Plus un prélèvement sera long, plus la concentration mesurée sera faible. Cette tendance, homogène à travers l'ensemble des agents chimiques, montre que pour une augmentation de la durée du prélèvement de 30 minutes, la concentration diminuera, en fonction de la strate, de 4 % à 5 % pour un prélèvement de longue durée et de 30 % à 40 % pour un prélèvement de courte durée. Ces différences ont déjà été identifiées lors de l'analyse des données de formaldéhyde dans COLCHIC [Lavoué et al., 2005] et lors de l'étude du trichloroéthylène sur les lieux de travail au Danemark [Raaschou-Nielsen, 2002]. Le niveau de concentration, plus de deux fois supérieur [1,85-2,05] pour les prélèvements de courte durée comparés aux prélèvements de longue durée, peut s'expliquer par une stratégie d'échantillonnage différente dans ces deux situations : en effet, un prélèvement de courte durée est généralement employé pour caractériser une tâche exposante alors qu'un prélèvement de longue durée couvre habituellement une

période représentative de l'exposition journalière du travailleur, pondérant ainsi les pics d'exposition auxquels il pourrait être soumis.

Une baisse globale des niveaux mesurés pour tous les agents chimiques [4 % - 10 %] est remarquée annuellement, quels que soient la période, la durée et le type de prélèvement. Cette tendance est moins marquée sur la période totale [4 % - 5 %] comparée à la période la plus récente [8 % - 10 %], mais est homogène pour les agents chimiques d'une même période. Ce constat de diminution des niveaux d'exposition au fil des années conforte celui déjà posé par plusieurs auteurs tels que Gomez [1997], Symanski et al. [1998, 2001], Okun et al., [2004] et Lavoué et al. [2005]. Cette baisse s'explique par une amélioration globale des conditions de travail des salariés entre 1987 ou 2002 et 2015 liée entre autres, aux actions de prévention telles que la formation et l'information des employeurs et des travailleurs.

L'équipement de protection individuelle est fortement lié à l'exposition dans le modèle. Lorsqu'il est adapté, c'est-à-dire que l'appareil de protection respiratoire porté par le travailleur permet de le protéger efficacement, les niveaux de concentration sont bien plus élevés [13 % - 66 %]. Cette tendance, globalement homogène pour tous les agents chimiques, insinue un biais sous-jacent dans les données et peut s'expliquer par le fait qu'un travailleur s'équipera d'un appareil de protection respiratoire dès lors que les niveaux d'exposition sont présumés élevés.

Codé à partir de 2002, le prédicteur type de procédé est significativement associé aux concentrations mesurées. En effet, quel que soit le type de prélèvement, un procédé semi-ouvert ou fermé présentera des niveaux plus faibles [4 % - 28 %] que s'il est ouvert. Cette tendance est cohérente puisqu'un procédé ouvert est plus émissif et exposera vraisemblablement plus un travailleur.

Globalement, lorsque l'exposition est permanente ou fréquente, les niveaux enregistrés sont plus importants que si elle est occasionnelle. Néanmoins, cette tendance ne se vérifie pas lors d'une utilisation occasionnelle de solvant en ambiance sur la période la plus récente. L'association de cette variable avec l'exposition est moins cohérente à travers toutes les données de COLCHIC.

L'origine de la demande d'intervention est faiblement associée aux concentrations

mesurées. Il est tout de même constaté que les niveaux d'exposition sont généralement plus faibles lorsque la démarche est à l'initiative de l'entreprise que lorsqu'elle est initiée par l'organisme de prévention ou le corps médical. Cette tendance n'est cependant pas confirmée pour les métaux sur la période 1987-2015, quel que soit le type de prélèvement. Ce constat général peut mettre en évidence un biais sur la représentativité des mesures par rapport à la population générale, lié au fait que les préventeurs ciblent les entreprises présentant des risques soupçonnés ou avérés. Cette conclusion peut être mise en parallèle de celles de la littérature [Froines et al., 1986 ; Froines et al., 1990 ; Gomez, 1997 ; Melville et Lippmann, 2001 ; Lavoué et al., 2008 ; Henn et al., 2011 ; Lavoué et al., 2013 ; Sarazin et al., 2016].

La présence d'une ventilation générale est corrélée à des concentrations plus élevées que lorsqu'il n'en existe pas. Cette observation identifiée par Lavoué et al. [2005] n'est pas confirmée lors de l'analyse spécifique des métaux ni lors de l'analyse sur la période 2002 à 2015. Cette variable est globalement peu associée aux niveaux comme l'avaient constaté Clerc et al. [2014].

Lorsque le motif de la demande d'intervention n'est pas lié à un risque d'exposition, les niveaux en individuel enregistrés dans COLCHIC, quelle que soit la période, sont plus bas [2 % - 14 %]. Cette tendance est analogue pour les mesures d'ambiance 2002-2015 (3 %). Elle est cependant inversée pour la période 1987-2015 (augmentation de 8 %), notamment pour les métaux (augmentation de 19 %). Les entreprises dont l'effectif est inférieur à 20 salariés présentent les niveaux les plus bas. Dans notre étude, nous n'avons pas observé d'effet marquant de la variable « Tranche d'effectif » bien que dans la littérature l'association entre la taille de l'entreprise et le niveau d'exposition soit plus prononcée [Gomez, 1997 ; Melville et Lippmann, 2001 ; Middendorf, 2004 ; Henn et al., 2011 ; Sarazin et al., 2016]. Aucune association entre les régions et les niveaux n'est observée dans nos analyses. Les régions présentent des variations importantes d'une période à une autre (par exemple, pour la région PACA, sur la période totale, les niveaux individuels sont 25 % plus importants comparés à la région de référence alors qu'ils le sont seulement de 5 % sur la période 2002-2015). Les stratégies d'échantillonnage réalisées par les LIC ne présentent pas de différences spécifiques suffisamment importantes pour expliquer ces variations interrégions. Ce constat avait déjà été observé par Lavoué et al [2005]. Ils

avaient alors conclu à une très faible corrélation interlaboratoires lors de leurs analyses de COLCHIC sur le formaldéhyde. Plus récemment, Sarazin et al. [2016] ont aussi observé des différences importantes entre les régions de l'OSHA, lors de l'évaluation de l'association entre les niveaux d'exposition enregistrés dans IMIS et les informations auxiliaires à travers 77 agents chimiques.

L'étude des Forest plots, associés à chaque catégorie, a permis d'observer l'homogénéité des effets à travers les agents chimiques afin de savoir si, pour un prédicteur, la tendance est globale à travers tous les agents chimiques et donc homogène pour toutes les données de COLCHIC ou si au contraire, le portrait est spécifique à certains agents chimiques et donc moins cohérent à travers les données de COLCHIC. D'une manière générale, les effets liés aux variables « Durée de prélèvement » et « Année » sont homogènes à travers tous les agents chimiques, quels que soient le type de prélèvement et la période considérée. Par exemple, pour un taux de réduction annuelle du niveau de concentration pour les mesures individuelles de longue durée de la période 2002-2015 à 8 %, la variation interagents chimiques est relativement réduite [- 1 % pour le chrome et 21 % pour le cobalt]. Pour d'autres variables descriptives, les effets à travers les agents chimiques sont plutôt hétérogènes, même si des tendances générales pour certaines catégories se démarquent. Prenons un premier exemple : de 1987 à 2015, l'effet lié à l'absence de ventilation générale présente des contrastes marqués à travers les agents chimiques. Comparé à la référence « Absence d'une ventilation générale mais présence d'un captage à la source », le méthacrylate de méthyle présente des niveaux individuels deux fois plus bas (0,47) avec un IC très étroit [0,36-0,63] ne contenant pas l'estimé moyen global. A contrario, pour le tétrahydrofurane, l'effet est inverse et présente des niveaux individuels deux fois plus importants que ceux de la référence (2,26) avec un IC plutôt large [1,28-3,99] qui ne contient pas non plus l'estimé moyen global. Pour le xylène, l'effet est quant à lui similaire comparé à la référence. Son IC très étroit [0,88-1,14] intègre l'estimé moyen global (1,06). Un deuxième exemple permet d'illustrer les différences entre les deux périodes étudiées. Dans le cas du port d'un équipement de protection adapté, comparé à la référence « Non adapté », les niveaux individuels pour le formaldéhyde sont stables d'une période à l'autre (2,01 de 1987 à 2015 et 1,98 de 2002 à 2015). A l'inverse, sur la période 1987-2015 l'équipement individuel adapté pour l'hexane

n'est pas associé au niveau (1,02) alors que sur 2002-2015 l'effet est plus marqué (1,17).

Ces analyses par modélisation statistique ont permis d'estimer l'effet sur les niveaux mesurés de 8 et 10 variables descriptives issues des informations enregistrées dans COLCHIC, respectivement pour les périodes 1987-2015 et 2002-2015. Quatre variables « Durée de prélèvement », « Année », « Equipement de protection individuelle » et « Type de procédé » ont une forte association sur les niveaux enregistrés à travers un grand nombre d'agents chimiques, quels que soient la période et le type de prélèvement. Trois autres variables « Fréquence d'exposition », « Origine de la demande » et « Ventilation générale » ont une association sur les niveaux d'exposition, mais leur portrait est moins cohérent à travers toutes les données de COLCHIC. Enfin, les trois dernières variables « Région », « Motif de la demande » et « Tranche d'effectif » ont, d'une manière générale, une faible association sur les niveaux mesurés. Pour ces 6 dernières variables, il est plus difficile de donner une interprétation globale.

Les implications des associations mises en évidence sur l'utilisation de COLCHIC et les orientations que dégagent nos résultats pour améliorer l'interprétation de COLCHIC sont discutées plus en détail dans le chapitre VII.

V.4.2. LIMITES DE L'ANALYSE

L'analyse de COLCHIC présente un portrait « global » des effets des variables descriptives sur les niveaux de concentration. Elle ne se veut pas spécifique à un agent chimique particulier. La méthodologie déployée, notamment pour construire les modèles, consiste à retenir les structures « Activité économique » et « Tendance temporelle et Durée de prélèvement » qui répondent aux critères qualité du plus grand nombre d'agents. C'est pour cette raison que l'approche entreprise, pour mener à bien ce travail, ne permet pas d'analyser spécifiquement chaque agent. Le choix des modèles finaux résulte donc d'une structure commune qui n'est pas spécifique. Il conviendra, lors de futurs travaux de modélisation, d'adapter le modèle au besoin de l'étude.

L'utilisation d'une structure complète, c'est-à-dire avec toutes les variables, est souvent déconseillée en modélisation. Cette approche peut en effet conduire à identifier des

effets observés qui sont liés à un phénomène de chance. Pour éviter cela, l'utilisation de procédures pour éliminer des modèles les variables peu prédictives est recommandée. Dans notre cas, les limites liées à l'approche retenue sont compensées par la Méta-analyse. En effet, même si des effets dus à la chance apparaissent pour des agents chimiques, ils se compensent pour donner un Méta coefficient global proche de 1.

L'agrégation de catégories de variables pour assurer un nombre suffisant de données dans chacune d'elles peut occasionner une perte de leur sens. Par exemple, le « Motif de la demande » est codé dans COLCHIC à partir de 17 catégories différentes. Celle portant sur le risque d'exposition représente une très large majorité de la codification. Compte tenu de la faible proportion de codification des 16 autres catégories, nous avons été amenés à les agglomérer dans une seule catégorie. Dans nos analyses cette variable compte seulement deux catégories.

V.5. CONCLUSION

Cette étude a permis de mettre en évidence des associations appuyées entre les niveaux de concentration enregistrés dans COLCHIC et quatre variables descriptives « Durée de prélèvement », « Equipement de protection individuelle », « Type de procédé » et « Année ». Globalement, des niveaux plus importants sont mesurés, pour les prélèvements de courte durée, lorsqu'un travailleur est équipé d'une protection individuelle adaptée ou qu'un procédé est ouvert.

Nos résultats suggèrent que le type de procédé mis en œuvre et les moyens de protection du travailleur influent sur les niveaux d'exposition enregistrés dans COLCHIC. Bien que les tendances observées à travers nos analyses ne peuvent pas être utilisées directement pour quantifier dans quelle mesure COLCHIC est représentative de la population générale, elles pourront être prises en compte lors de l'utilisation des données à des fins d'évaluation de l'exposition. Malgré certaines limites liées à la codification de l'information et à la représentativité des mesures par rapport à la population générale, COLCHIC reste une source importante d'informations sur les expositions professionnelles dans le cadre d'évaluations des risques et des expositions et en épidémiologie.

CHAPITRE VI. MODELISATION COMPARATIVE DES BASES DE DONNÉES COLCHIC ET SCOLA

VI.1. INTRODUCTION

Ce chapitre décrit tout d'abord la relation entre les variables et les niveaux d'exposition enregistrés dans COLCHIC et SCOLA pour la période 2007-2015 puis il présente leur comparaison sur cette même période. Il permet de finaliser le premier objectif de la thèse et de répondre au second.

La coexistence, depuis le déploiement de SCOLA en 2007, de deux bases de données françaises portant sur des expositions aux produits chimiques est une situation atypique sur le plan international. L'originalité repose sur le fait que ces deux bases enregistrent des niveaux d'exposition mesurés avec des objectifs différents (prévention et réglementation) pour un tissu industriel identique (Chapitre III).

Au regard des travaux du chapitre IV portant sur l'analyse descriptive de COLCHIC et de SCOLA pour la période commune 2007-2012, il a été mis en évidence des différences systématiques entre ces deux BDEP. Premièrement, les objectifs des campagnes de mesures divergent avec entre autres pour conséquence une certaine disparité dans les secteurs d'activité investigués. Deuxièmement, les niveaux d'exposition enregistrés sont généralement plus importants dans COLCHIC que dans SCOLA. Cette tendance s'est vérifiée lors de l'analyse des 153 entreprises communes à COLCHIC et SCOLA sur la période 2007-2012. Troisièmement, l'aspect réglementaire de SCOLA engendre un enregistrement quasiment exclusif de mesures individuelles. Dans ces circonstances, il est paru important d'explorer ces différences plus en profondeur au moyen de la modélisation statistique.

Ainsi, à travers ce travail, nous avons cherché dans un premier temps, à identifier si les associations entre les expositions et les variables enregistrées dans COLCHIC et SCOLA étaient similaires puis dans un deuxième temps, à évaluer les différences absolues au travers d'une analyse regroupant les données de COLCHIC et SCOLA.

VI.2. MÉTHODE

Dans cette section, nous présentons succinctement les informations majeures issues du chapitre III « Méthodes » auxquelles viennent s'ajouter, de manière plus détaillée, les spécificités de ces analyses par modélisation.

VI.2.1. PRÉPARATION DES DONNÉES

Tout d'abord, pour chacune des BDEP, les résultats des mesures prélevées en individuel sur une durée comprise entre 1 et 720 minutes et exprimées en unité d'air ont été sélectionnées pour la période 2007 à 2015. Les agents chimiques comptant plus de 1 000 mesures par BDEP sont conservés. Les mesures d'amiante sont écartées compte tenu d'une part, des évolutions analytiques courant 2012 rendant caduc une large majorité de données et d'autre part, du faible nombre de données dans COLCHIC. Une LoQ empirique est attribuée aux résultats enregistrés comme « Non détecté ».

Après avoir préparé les jeux de données propres à COLCHIC et SCOLA, un troisième jeu de données, appelé « BDEP commune », fusionne les mesures issues des deux BDEP pour des agents chimiques communs. Seules les variables communes aux deux BDEP sont conservées. L'architecture fonctionnelle de COLCHIC et SCOLA étant très proche, il n'a pas été nécessaire de procéder à un recodage des variables étudiées.

VI.2.2. VARIABLES EXPLOITÉES

Dans l'analyse de SCOLA et de la BDEP commune, nous avons inclus 13 variables « Année », « Entreprise », « Durée de prélèvement », « Protection individuelle », « Effectif », « Equipement de protection collective », « Tâche », « Métier », « Secteur d'activité », « Région », « Type d'échantillonnage », « Durée stratifiée » et « Intervention » auxquelles s'additionnent 4 autres variables « Origine de la demande », « Motif de la demande », « Type de procédé » et « Fréquence d'exposition » pour COLCHIC.

Afin de différencier la provenance des données lors de l'analyse comparative, une variable dichotomique « Source » composée de « COLCHIC » et « SCOLA » est créée.

VI.2.3. MODÉLISATION STATISTIQUE POUR CHAQUE AGENT

L'intérêt de cette étude consiste à identifier les différences et les similitudes entre COLCHIC et SCOLA. Pour cela, nous nous sommes appuyés sur la méthodologie décrite dans le chapitre III.

VI.2.3.1. Analyse séparée de COLCHIC et SCOLA

Les modèles Tobit ont ainsi été construits en incluant toutes les variables d'intérêt propres à chaque BDEP. Chronologiquement, les structures « Activité économique » et « Tendance temporelle et durée de prélèvement » ont été définies les unes après les autres à partir des critères de qualité AIC et BIC pour définir leur meilleure combinaison. La qualité du modèle final est mesurée pour chaque agent chimique avec les coefficients de détermination (R^2).

VI.2.3.2. Analyse comparative de COLCHIC et SCOLA

Le mode de construction des modèles pour l'analyse comparative de COLCHIC et SCOLA est identique pour celui de leurs analyses séparées. Dans un premier temps, nous avons ajouté simplement la variable « Source » afin d'estimer la différence entre COLCHIC et SCOLA en tenant compte de toutes les associations. Néanmoins, ce modèle ne permet pas de modéliser la possibilité d'association différente entre la « Source » et les autres variables. Nous avons donc construit, dans un deuxième temps, un autre modèle qui inclut une quatrième structure portant sur l'interaction entre la « Source » et les autres variables. Plusieurs modèles, combinant tout d'abord la « Source » avec chacune des variables d'intérêt et ensuite la « Source » avec plusieurs variables, ont été testés. Le choix final s'est porté sur le modèle obtenant les meilleurs AIC et BIC. En complément, pour simplifier leur exploitation, les coefficients de SCOLA issus de l'analyse comparative prenant en compte les interactions ont été pondérés pour que la référence soit à 1,00.

Afin d'évaluer une différence moyenne entre COLCHIC et SCOLA, nous avons créé 4 scénarios pour refléter les extrêmes des variables les plus influentes dans les analyses séparées : 2007/2015 et 30/240 minutes. Pour les autres variables, nous avons considéré

un effet moyen à travers les catégories de chaque variable, excluant les catégories « Non documentées » par exemple.

$$C = \exp(\alpha + 1/13 \sum_1^{13} \beta_{i \text{ Région}} + 1/4 \sum_1^4 \beta_{i \text{ Ventilation}} + 1/4 \sum_1^4 \beta_{i \text{ Effectif}} + 1/2 \beta_{\text{Prot adaptée}} + x_1 \beta_{\text{Année}} + x_2 \beta_{\text{Durée prlvt}} + x_3 \beta_{\text{Type échantillon}} + x_4 \beta_{\text{Année:Type échantillon}} + x_5 \beta_{\text{Durée prlvt:Type échantillon}})$$

α : Intercept du modèle

β_i : Coefficient de chaque catégorie de variable

x : Facteur multiplicatif. Par exemple, pour faire l'estimation de la pente de la durée pour un prélèvement de 240 minutes, x_2 sera égal à 240.

Les autres variables ont été interprétées comme décrites dans le chapitre III, à la différence que leur effet est estimé séparément, pour les variables COLCHIC et SCOLA, à l'aide des coefficients d'interaction.

Afin de répondre à certaines interrogations, plusieurs analyses de sensibilité ont été réalisées. Comme décrit dans le chapitre V, une première analyse a porté sur l'impact de la structure aléatoire sur les coefficients du modèle. Une deuxième analyse a permis d'étudier l'impact sur les associations lorsque les données sélectionnées portent uniquement sur les secteurs d'activités communs à COLCHIC et SCOLA. Afin de confirmer qu'en conservant tous les secteurs d'activité aucun artefact ne serait introduit, nous avons contrôlé que les résultats, en ne conservant que les secteurs d'activité communs aux deux BDEP, étaient similaires.

Enfin, pour s'assurer que les associations évaluées soient comparables, les mêmes structures pour l'« Activité économique » et la « Tendance temporelle et durée de prélèvement » doivent être choisies pour les 3 jeux de données.

VI.2.4. MÉTA ANALYSE

La Méta analyse est l'approche retenue pour synthétiser les coefficients de la modélisation. Un coefficient global, synthétisant la tendance générale des coefficients de chaque agent chimique, est calculé. Il est exprimé sous la forme d'un RIE pour les

variables de type catégorie et correspond à une pente moyenne à tous les agents chimiques pour les variables continues. En complément, une Méta analyse stratifiée pour les agents chimiques rattachés à la famille des solvants a été réalisée lorsque le nombre de substances était suffisant.

L'utilisation de Forest plots a permis de représenter graphiquement la variabilité interagents chimiques.

VI.3. RÉSULTATS

VI.3.1. ANALYSE DESCRIPTIVE

Sur la période 2007–2015, les extractions de COLCHIC et SCOLA comptent respectivement 290 132 et 429 104 données correspondant à 9 727 et 124 316 interventions dans 6 795 et 19 813 établissements. Après avoir appliqué les critères de sélection, nous avons éliminé 59 % (n = 170 466) des données de COLCHIC, dont 37 % de mesures d’ambiance et 49 % (n = 208 937) de celles de SCOLA, dont 45 % de mesures d’amiante. La concaténation des données pour la BDEP commune, restreinte aux 15 agents chimiques communs, comptabilise 239 968 données (25 % COLCHIC et 75 % SCOLA) (Figure VI.1).

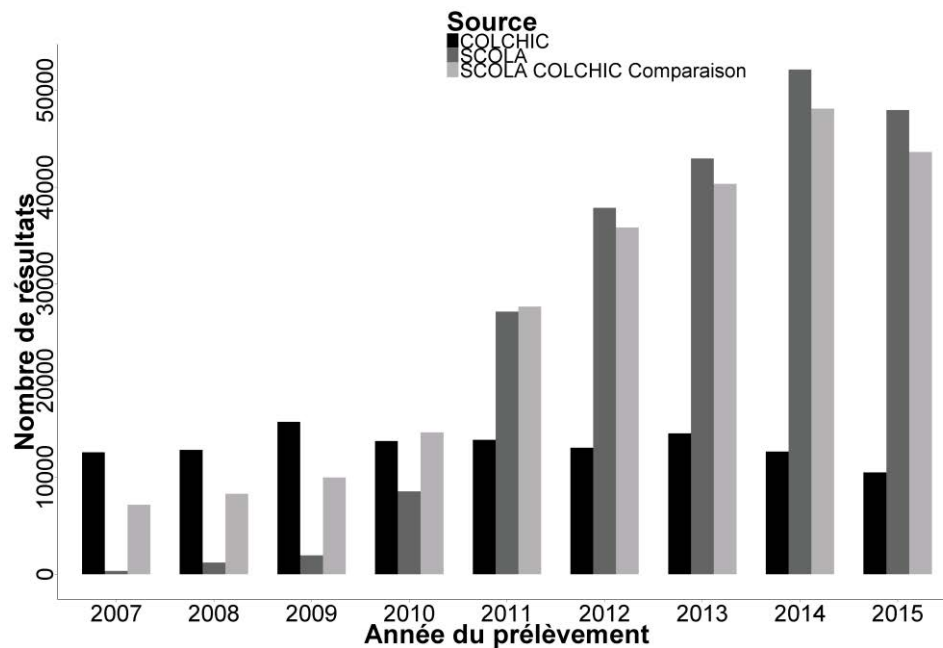


Figure VI.1 Distribution annuelle du nombre de mesures individuelles dans COLCHIC, SCOLA et la BDEP commune sur la période 2007-2015

Pour chacune des variables incluses dans les modèles, la distribution du nombre de données stratifiées selon leur source est synthétisée Tableau VI.1. COLCHIC, SCOLA et la BDEP commune contiennent respectivement 15 %, 18 % et 14 % de mesures de courte durée.

Tableau VI.1 Distribution du nombre de données individuelles (%) par catégorie de variables incluses dans les modèles pour la période 2007-2015 pour COLCHIC, SCOLA et la BDEP commune (Comp. C&S)

Variable	Description / Type	COLCHIC	SCOLA	Comp. C&S
Effets fixes				
Source	Provenance des données / Dichotomique (2 catégories)			
	COLCHIC	-	-	59243 (25)
	SCOLA	-	-	176485 (75)
Région	Région où les prélèvements ont été effectués / Nominal (13 catégories)			
	Poitou Charentes Aquitaine Limousin (PCLA)	10617 (8,9)	17482 (7.8)	19824 (8.4)
	Auvergne Rhône Alpes (ARA)	13364 (11)	32759 (15)	32234 (14)
	Bourgogne Franche Comté (Bourgo_comte)	9281 (7,8)	11119 (5.1)	13998 (5.9)
	Bretagne	4815 (4)	8754 (4)	9677 (4.1)
	Champagne Ardenne Lorraine Alsace (CALA)	8673 (7,2)	20471 (9.2)	21457 (9.1)
	Centre	10140 (8,5)	9161 (4.2)	12491 (5.3)
	Département d'Outre-Mer et Corse (DOM Corse)	6456 (5,4)	3999 (1.8)	5964 (2.5)
	Ile de France (IDF)	15390 (13)	25992 (12)	27695 (12)
	Midi-Pyrénées Languedoc Roussillon (MPLR)	18725 (16)	11133 (5.1)	19293 (8.2)
	Nord Pas-de-Calais Picardie (Nord)	4746 (4)	26937 (12)	22967 (9.7)
	Normandie	6691 (5,6)	24234 (11)	21407 (9.1)
	Pays de la Loire (P_Loire)	7239 (6)	13896 (6.3)	16071 (6.8)
	Provence Alpes Côte d'Azur (PACA)	3529 (2,9)	14230 (6.5)	12650 (5.4)
Motif de la demande	Information sur le motif de la demande d'intervention / Dichotomique (2 catégories)			
	Risque d'exposition	99220 (83)	-	-

Variable	Description / Type	COLCHIC	SCOLA	Comp. C&S
Origine de la demande	Autre raison	20446 (17)	-	-
	Information sur l'origine de la demande d'intervention / Nominal (4 catégories)			
	Organisme prévention	72619 (61)	-	-
	Entreprise	20868 (17)	-	-
	Corps médical	23071 (19)	-	-
	Autre origine	3108 (3)	-	-
Ventilation globale	Description de la ventilation générale sur le lieu de travail / Nominal (5 catégories)			
	Non mais présence d'un captage localisé	41830 (35)	61364 (28)	74502 (32)
	Oui et présence d'un captage localisé	16712 (14)	36863 (17)	35601 (15)
	Oui	14040 (12)	22376 (10)	23378 (9)
	Non	39851 (33)	72736 (33)	76852 (33)
	Information non relevée	7233 (6)	26828 (12)	25395 (11)
Tranches d'effectif	Nombre de travailleurs dans l'établissement visité / Nominal (5 catégories)			
	1-19 salariés	26858 (22)	39172 (18)	49946 (21)
	20-49 salariés	29982 (25)	27845 (13)	38582 (16)
	50-199 salariés	36810 (31)	60727 (28)	67118 (28)
	200 salariés ou plus	15188 (13)	92423 (42)	74718 (32)
	Information non relevée	10828 (9)	-	5364 (3)
Equipement de protection individuelle	Description de l'appareil de protection respiratoire du travailleur / Nominal (3 catégories)			

Variable	Description / Type	COLCHIC	SCOLA	Comp. C&S
	Non adapté	83066 (69)	141609 (64)	152319 (65)
	Adapté	30431 (25)	58170 (26)	64519 (27)
	Information non relevée	6169 (6)	20388 (10)	18890 (8)
Type de procédé	Description du type de procédé à l'origine de l'exposition / Nominal (4 catégories)			
	Ouvert	98785 (83)	-	-
	Semi-ouvert	11419 (9)	-	-
	Fermé	7460 (6)	-	-
	Autre type de procédé	2002 (2)	-	-
Fréquence d'exposition	Fréquence d'exposition du travailleur / Nominal (4 catégories)			
	Permanent	71251 (60)	-	-
	Fréquent	33157 (28)	-	-
	Occasionnel	13537 (11)	-	-
	Autre fréquence d'exposition	1721 (1)	-	-
Durée d'échantillonnage	Durée du prélèvement / Dichotomique (2 catégories)			
	Longue durée	101472 (85)	181059 (82)	202820 (86)
	Courte durée	18194 (15)	39108 (18)	32908 (14)
Année	Année du prélèvement / Continue			
Durée de prélèvement	Durée du prélèvement / Continue			
Effets aléatoires				
Intervention	Triptyque Année / Laboratoire / Dossier COLCHIC ou SCOLA			

L'Annexe VI.1 présente la distribution du nombre de mesures par agent chimique comptant plus de 1 000 résultats sur 2007-2015, stratifiée selon les jeux de données de COLCHIC, SCOLA et la BDEP commune. Pour ces trois jeux de données, nous recensons respectivement 39, 35 et 17 agents chimiques, dont 46 %, 60 % et 65 % de solvants, comptant plus de 1 000 mesures individuelles de 2007 à 2015. Les 4 agents chimiques les plus mesurés dans COLCHIC sont les poussières inhalables ($n = 12\,107$), l'acétone ($n = 7\,117$), le fer ($n = 6\,302$) et le styrène ($n = 5\,150$). Pour SCOLA, il s'agit des poussières de bois ($n = 37\,385$), des poussières alvéolaires ($n = 22\,383$), du quartz ($n = 21\,176$) et de la cristobalite ($n = 18\,719$). Et enfin, pour la comparaison des 2 BDEP, il s'agit des poussières de bois ($n = 42\,193$), des poussières alvéolaires ($n = 25\,299$), du quartz ($n = 23\,774$) et des poussières inhalables ($n = 22\,386$). La proportion médiane de données inférieures à la LoQ est respectivement pour COLCHIC, SCOLA et la BDEP commune de 29 % [2 - 94], 54 % [19 - 94] et 45 % [11 - 93] pour la période 2007-2015.

VI.3.2. STRUCTURES DES MODÈLES

Concernant la multicolinéarité, tous les couples de variables indépendantes ont une corrélation inférieure à 0,7. Ces conclusions ont permis de définir les variables à tester pour les modèles de chaque jeu de données.

Tout d'abord, le choix de la structure « Aléatoire » s'est porté sur la variabilité liée aux interventions, au détriment de celle liée à l'établissement (Chapitre III).

Puis, nous avons choisi la structure « Activité économique » et enfin la combinaison de la structure « Tendances temporelles et Durée de prélèvement ». Les Tableaux 2 et 3 présentent, pour chaque structure testée, stratifiée par source, la proportion entre le nombre de fois où l'AIC et le BIC sont les plus bas pour un agent chimique et le nombre total d'agents. Globalement, le message entre les deux critères qualité ne permet pas de conclure aisément sur le choix de la combinaison. Pour la structure « Activité économique », le modèle avec le secteur d'activité est retenu compte tenu de la quasi-unanimité des résultats du BIC pour les 3 sources, bien que le message soit plus variable pour le critère AIC. Bien que pour la structure « Tendances temporelles et Durée de prélèvement » aucune situation majoritaire ne se dégage pour COLCHIC et SCOLA, la

combinaison (9) est prépondérante pour la BDEP commune. Le choix s'est donc porté sur cette dernière pour les 3 jeux de données étudiés.

Tableau VI.2 Choix de la structure « Activité économique »

Combinaison de structure « Activité économique »	AIC			BIC		
	COLCHIC ⁽¹⁾	SCOLA ⁽¹⁾	Comp. C&S ⁽¹⁾	COLCHIC ⁽¹⁾	SCOLA ⁽¹⁾	Comp. C&S ⁽¹⁾
SECTEUR ACTIVITE	33	13	24	97	94	100
TACHE	28	23	12	0	0	0
METIER	39	60	65	3	6	0

⁽¹⁾ Proportion (%) du nombre de fois où l'AIC ou le BIC est le plus bas pour un agent chimique par rapport au nombre total d'agents chimiques évalués

Tableau VI.3 Choix de la structure « Tendances temporelles et durée de prélèvement »

Combinaison de structure « Tendances temporelles et durée de prélvt » ⁽¹⁾	AIC			BIC		
	COLCHIC ⁽²⁾	SCOLA ⁽²⁾	Comp. C&S ⁽²⁾	COLCHIC ⁽²⁾	SCOLA ⁽²⁾	Comp. C&S ⁽²⁾
(1)	3	3	0	11	0	0
(2)	11	10	0	11	5	0
(3)	9	29	25	0	9	6
(4)	3	0	0	9	18	6
(5)	0	0	0	0	5	0
(6)	14	6	0	29	27	25
(7)	17	19	6	11	23	25
(8)	23	19	13	23	14	25
(9)	20	13	56	6	0	13

⁽¹⁾ Structures « Tendances temporelles et durée de prélèvement » présentées au paragraphe 6.3 du chapitre « Méthodes »

⁽²⁾ Proportion (%) du nombre de fois où l'AIC ou le BIC est le plus bas pour un agent chimique par rapport au nombre total d'agents chimiques évalués

Le jeu de données utilisé pour comparer les deux BDEP intègre la variable « Source » permettant ainsi d'identifier l'origine de chaque mesure. L'interaction de cette variable avec les autres variables a amélioré la qualité des modèles selon les critères AIC et BIC et a donc été ajoutée au modèle pour l'analyse de la comparaison de COLCHIC et SCOLA.

VI.3.3. MODÉLISATION DES BASES DE DONNÉES D'EXPOSITION PROFESSIONNELLE

Dans ce chapitre, sont exposés les résultats de l'analyse séparée de COLCHIC et de SCOLA, puis ceux de la modélisation comparative de ces deux BDEP.

VI.3.3.1. Modélisation séparée de COLCHIC et SCOLA

La proportion de la variabilité des niveaux d'exposition expliquée par les effets fixes à travers l'ensemble des agents chimiques analysés est présentée pour les 3 jeux de données dans le Tableau VI.4.

Tableau VI.4 Coefficient de détermination par période et par type de prélèvement

Jeu de données	Coefficient de détermination (R^2)		
	Min (AC)	Médiane	Max (AC)
COLCHIC	13 % (Poussières de bois)	32 %	49 % (Titane)
SCOLA	15 % (Poussières de bois)	41 %	56 % (Cyclohexane)
Comp. C&S	18 % (Poussières de bois)	38 %	51 % (Ethylbenzène)

(AC) Agent chimique correspondant au coefficient de détermination

Le Tableau VI.5 recense les résultats des analyses séparées de COLCHIC et SCOLA. Les niveaux mesurés pour les prélèvements de longue durée sont globalement plus de deux fois plus faibles comparés à ceux de courte durée [0,39 - 0,62]. La tendance annuelle de l'exposition dans COLCHIC est à la baisse, quelle que soit la durée d'échantillonnage (7 % par année). Ce constat est vérifié pour 14/17 et 29/33 agents chimiques respectivement pour les prélèvements de courte et de longue durée. A l'inverse, la tendance dans SCOLA est à la hausse (5 % par année) pour les prélèvements de courte durée et nulle pour ceux de longue durée. Cette augmentation, pour les mesures de courte durée, est constatée pour une majorité d'agents chimiques (7/12). D'une région à une autre et quelle que soit la BDEP, les niveaux mesurés fluctuent. Lorsque le motif de la demande est lié à un risque d'exposition, les niveaux rencontrés dans COLCHIC sont plus importants (14 %). Si un

organisme de prévention est à l'origine de la demande, les niveaux mesurés dans COLCHIC sont, d'une manière générale, sensiblement plus importants [2 % - 6 %]. En présence d'une ventilation générale, les niveaux dans COLCHIC sont globalement plus élevés. Dans SCOLA, la présence d'un captage localisé seul (sans ventilation générale) est la situation la plus exposante [19 % - 21 %]. Globalement, les niveaux sont faiblement associés aux tranches d'effectif dans COLCHIC. Au contraire, dans SCOLA les concentrations sont inversement proportionnelles à la taille de l'entreprise. Par exemple, une entreprise de moins de 20 salariés expose plus ses travailleurs à un agent chimique qu'une entreprise de plus de 200 salariés (35 %). Un équipement de protection individuelle adapté est associé à une forte concentration comparée aux situations où la protection est absente ou non adaptée (facteur 1,8 et 2,5 pour COLCHIC et SCOLA). Cette tendance est homogène pour 31/33 et 20/20 agents respectivement dans COLCHIC et SCOLA. Les effets du « Type de procédé » et de la « Fréquence d'exposition » sont semblables aux analyses de COLCHIC faites sur la période plus large. Ainsi, un procédé ouvert est lié à des niveaux plus importants comparés à un procédé semi-ouvert (26 %) ou fermé (29 %) et une exposition permanente ou fréquente engendre des niveaux plus importants que si elle est occasionnelle (16 % - 20 %).

Nous avons stratifié les résultats par famille pour évaluer des tendances dans COLCHIC et SCOLA, mais globalement nous n'avons rien identifié. Les résultats détaillés sont présentés Annexes VI.2 et VI.3.

Les analyses de sensibilité réalisées dans le but de mesurer l'impact de la structure aléatoire ont montré des résultats similaires à l'analyse principale. Les résultats détaillés sont présentés Annexes VI.2 et VI.3.

Tableau VI.5 Synthèse des indices d'exposition relative (RIE) et de leur intervalle de confiance à 95 % (IC 95 %) par Méta analyse pour COLCHIC et SCOLA sur la période 2007-2015

Variable	COLCHIC	SCOLA
Catégorie	N=33 RIE (IC 95 %)	N=20 RIE (IC 95 %)
Durée du prélèvement		
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :		
Courte durée (< 60 minutes) ⁽¹⁾	34 % [27-41]	37 % [21-50]
Longue durée (≥ 60 minutes)	5 % [3-6]	4 % [3-6]
Ratio Conc(240)/Conc(30) en 2007⁽¹⁾	0,51 (0,41;0,62)	0,49 (0,29;0,83)
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,42 (0,37;0,48)	0,39 (0,32;0,48)
Année du prélèvement		
Taux de réduction annuelle de l'exposition :		
Courte durée (< 60 minutes) ⁽¹⁾	7 % [3-11]	-5 % [-14-3]
Longue durée (≥ 60 minutes)	7 % [5-9]	0 % [-4-4]
Région		
PCLA	1,00 (Réf.)	1,00 (Réf.)
ARA	1,19 (0,98;1,45)	1,09 (0,93;1,28)
Bourgo_comte	0,90 (0,75;1,06)	1,07 (0,89;1,28)
Bretagne	0,87 (0,72;1,05)	1,27 (1,00;1,61)
CALA	1,04 (0,90;1,21)	1,00 (0,84;1,19)
Centre	0,98 (0,85;1,14)	0,91 (0,73;1,15)
DOM Corse	0,96 (0,79;1,17)	1,22 (0,84;1,77)
IDF	1,04 (0,87;1,25)	1,04 (0,85;1,27)
MPLR	1,01 (0,85;1,21)	1,20 (1,02;1,42)
Nord	1,32 (1,02;1,72)	1,06 (0,88;1,29)
Normandie	1,15 (0,93;1,41)	0,99 (0,79;1,25)
P_Loire	0,86 (0,74;1,00)	1,18 (0,99;1,41)
PACA	0,92 (0,71;1,18)	1,12 (0,95;1,31)
Motif de la demande		
Risque d'exposition	1,00 (Réf.)	-
Autre raison	0,86 (0,73;1,01)	-
Origine de la demande		
Organisme de prévention	1,00 (Réf.)	-
Entreprise	0,97 (0,84;1,12)	-
Corps médical	0,98 (0,83;1,15)	-
Autre origine	0,94 (0,73;1,21)	-

Variable	COLCHIC	SCOLA
Catégorie	N=33 RIE (IC 95 %)	N=20 RIE (IC 95 %)
Ventilation générale		
Non mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,02 (0,91;1,15)	0,81 (0,72;0,91)
Oui	1,08 (0,94;1,23)	0,81 (0,70;0,93)
Non	0,95 (0,86;1,04)	0,80 (0,69;0,91)
Information non relevée	0,84 (0,77;0,92)	0,79 (0,70;0,89)
Tranches d'effectif		
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,10 (0,99;1,22)	0,88 (0,78;0,98)
50-199 salariés	1,01 (0,89;1,15)	0,85 (0,73;0,99)
200 salariés ou plus	1,08 (0,91;1,28)	0,65 (0,53;0,79)
Information non relevée	1,07 (0,94;1,22)	-
Equipement de protection individuelle		
Non adapté	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,82 (1,63;2,03)	2,51 (2,05;3,06)
Information non relevée	1,04 (0,92;1,16)	1,03 (0,91;1,17)
Type de procédé		
Ouvert	1,00 (Réf.)	-
Semi-ouvert	0,74 (0,65;0,84)	-
Fermé	0,71 (0,62;0,81)	-
Autre type de procédé	0,80 (0,59;1,08)	-
Fréquence d'exposition		
Permanent	1,00 (Réf.)	-
Fréquent	0,96 (0,89;1,04)	-
Occasionnel	0,80 (0,75;0,84)	-
Autre fréquence d'exposition	0,66 (0,52;0,85)	-

⁽¹⁾Données calculées avec 17 et 12 agents chimiques respectivement pour COLCHIC et SCOLA

VI.3.3.2. Modélisation de la comparaison de COLCHIC et SCOLA

Tout d'abord, l'analyse du modèle simple, sans prise en compte des interactions, nous a permis d'observer que, d'une manière générale, les expositions dans SCOLA sont plus faibles que celles dans COLCHIC d'un facteur 0,69 (Annexe VI.4).

Puis, lors de l'analyse du jeu de données commun avec les interactions, nous avons pu estimer l'effet des variables séparément pour COLCHIC et SCOLA (Tableau VI.6). Les associations mesurées par la modélisation comparative confirment les observations de la

modélisation séparée. elles mettent en avant des similitudes avec celles obtenues lors de leurs analyses séparées. Par exemple, concernant la durée d'échantillonnage et le taux de réduction annuelle de l'exposition, les résultats sont équivalents : mesures de longue durée plus faibles que les prélèvements de courte durée [0,38 – 0,50]. Il en est de même pour la tendance annuelle, toujours à la baisse pour COLCHIC (9 %) et à la hausse (4 %) ou stable pour SCOLA respectivement pour les mesures de courte durée et de longue durée. Les régions ne présentent pas de tendance particulière. La présence d'une ventilation générale associée ou non à un captage localisé dans COLCHIC et la présence seule d'un captage à la source dans SCOLA coïncident avec des niveaux plus importants respectivement [4 % - 8 %] et [12 % - 25 %] comparés à la référence. Excepté dans l'analyse séparée de SCOLA, les concentrations sont faiblement associées aux tranches d'effectif dans les autres jeux de données : concentrations plus élevées dans les entreprises de 1 à 19 salariés comparées aux entreprises avec effectif plus important [6 % - 29 %]. Enfin, l'emploi d'équipements de protection individuelle adaptés est associé à des niveaux plus élevés que lorsqu'ils ne sont pas adaptés : facteur 1,94 pour COLCHIC et 2,52 pour SCOLA (Tableau VI.7).

Tableau VI.6 Synthèse des indices d'exposition relative (RIE) et de leur intervalle de confiance à 95 % (IC 95 %) par Méta analyse pour la BDEP commune avec les interactions entre la « Source » et les autres variables sur la période 2007-2015

Variable Catégorie	Comp C&S avec interactions	
	COLCHIC N=15 RIE (IC 95 %)	SCOLA N=15 RIE (IC 95 %)
Durée du prélèvement		
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :		
Courte durée (< 60 minutes) ⁽¹⁾	36 % [30;40]	36 % [0;55]
Longue durée (≥ 60 minutes)	4 % [2-5]	4 % [3-5]
Ratio Conc(240)/Conc(30) en 2007⁽¹⁾	0,62 (0,54;0,71)	0,44 (0,21;0,91)
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,50 (0,47;0,53)	0,38 (0,20;0,71)
Année du prélèvement		
Taux de réduction annuelle de l'exposition :		
Courte durée (< 60 minutes) ⁽¹⁾	9 % [4-14]	-4 % [-9-1]
Longue durée (≥ 60 minutes)	9 % [5-12]	0 % [-4-5]

Variable Catégorie	Comp C&S avec interactions	
	COLCHIC N=15 RIE (IC 95 %)	SCOLA N=15 RIE (IC 95 %)
Région		
PCLA	1,00 (Réf.)	0,48 (Réf.)
ARA	0,90 (0,77;1,04)	0,56 (0,36;0,86)
Bourgo_comte	0,91 (0,78;1,05)	0,60 (0,38;0,96)
Bretagne	0,87 (0,72;1,05)	0,57 (0,37;0,87)
CALA	1,14 (0,91;1,43)	0,54 (0,33;0,86)
Centre	1,04 (0,82;1,33)	0,42 (0,27;0,65)
DOM Corse	0,94 (0,72;1,24)	0,66 (0,44;0,99)
IDF	0,91 (0,79;1,05)	0,52 (0,32;0,83)
MPLR	0,93 (0,77;1,11)	0,49 (0,31;0,77)
Nord	1,26 (0,83;1,91)	0,56 (0,35;0,92)
Normandie	0,97 (0,81;1,16)	0,51 (0,32;0,80)
P_Loire	0,86 (0,71;1,04)	0,60 (0,41;0,89)
PACA	0,84 (0,61;1,16)	0,57 (0,35;0,91)
Ventilation générale		
Non mais présence d'un captage localisé	1,00 (Réf.)	0,48 (Réf.)
Oui et présence d'un captage localisé	1,08 (0,95;1,22)	0,36 (0,23;0,55)
Oui	1,04 (0,95;1,13)	0,38 (0,24;0,59)
Non	0,94 (0,84;1,05)	0,37 (0,24;0,58)
Information non relevée	0,80 (0,68;0,95)	0,42 (0,27;0,65)
Tranches d'effectif		
1-19 salariés	1,00 (Réf.)	0,48 (Réf.)
20-49 salariés	1,06 (0,97;1,15)	0,45 (0,29;0,70)
50-199 salariés	0,91 (0,82;1,02)	0,41 (0,26;0,65)
200 salariés ou plus	0,98 (0,76;1,27)	0,34 (0,21;0,53)
Information non relevée	1,17 (0,95;1,43)	-
Equipe ment de protection individuelle		
Non adapté	1,00 (Réf.)	0,48 (Réf.)
Adapté	1,94 (1,71;2,20)	1,21 (0,79;1,86)
Information non relevée	0,97 (0,86;1,10)	0,52 (0,35;0,76)

⁽¹⁾Données calculées avec 9 agents chimiques pour la BDEP commune

La comparaison des coefficients issus des analyses séparées de COLCHIC et SCOLA avec ceux du jeu de données commun est synthétisée dans le Tableau VI.7 et montre des similitudes dans leurs résultats.

Tableau VI.7 Synthèse des indices d'exposition relative (RIE) par Méta analyse pour COLCHIC, SCOLA et la BDEP commune avec les interactions entre la « Source » et les autres variables sur la période 2007-2015

Variable	COLCHIC	SCOLA	Comp C&S avec interactions	
Catégorie	N=33 RIE	N=20 RIE	COLCHIC N=15 RIE	SCOLA N=15 RIE ⁽²⁾
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	34 %	37 %	36 %	36 %
Longue durée (≥ 60 minutes)	5 %	4 %	4 %	4 %
Ratio Conc(240)/Conc(30) en 2007⁽¹⁾	0,51	0,49	0,62	0,44
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,42	0,39	0,50	0,38
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	7 %	-5 %	9 %	-4 %
Longue durée (≥ 60 minutes)	7 %	0 %	9 %	0 %
Région				
PCLA	1,00	1,00	1,00	1,00
ARA	1,19	1,09	0,90	1,17
Bourgo_comte	0,90	1,07	0,91	1,25
Bretagne	0,87	1,27	0,87	1,19
CALA	1,04	1,00	1,14	1,13
Centre	0,98	0,91	1,04	0,88
DOM Corse	0,96	1,22	0,94	1,38
IDF	1,04	1,04	0,91	1,08
MPLR	1,01	1,20	0,93	1,02
Nord	1,32	1,06	1,26	1,17
Normandie	1,15	0,99	0,97	1,06
P_Loire	0,86	1,18	0,86	1,25
PACA	0,92	1,12	0,84	1,19
Ventilation générale				
Non mais présence d'un captage localisé	1,00	1,00	1,00	1,00
Oui et présence d'un captage localisé	1,02	0,81	1,08	0,75
Oui	1,08	0,81	1,04	0,79
Non	0,95	0,80	0,94	0,77
Information non relevée	0,84	0,79	0,80	0,88

Variable	COLCHIC	SCOLA	Comp C&S avec interactions	
Catégorie	N=33 RIE	N=20 RIE	COLCHIC N=15 RIE	SCOLA N=15 RIE ⁽²⁾
Tranches d'effectif				
1-19 salariés	1,00	1,00	1,00	1,00
20-49 salariés	1,10	0,88	1,06	0,94
50-199 salariés	1,01	0,85	0,91	0,85
200 salariés ou plus	1,08	0,65	0,98	0,71
Information non relevée	1,07	-	1,17	-
Equipement de protection individuelle				
Non adapté	1,00	1,00	1,00	1,00
Adapté	1,82	2,51	1,94	2,52
Information non relevée	1,04	1,03	0,97	1,08

⁽¹⁾Données calculées avec 9 agents chimiques pour la BDEP commune

⁽²⁾Coefficients de SCOLA pondérés

Afin d'évaluer une différence moyenne entre COLCHIC et SCOLA, nous avons calculé le ratio SCOLA/COLCHIC pour chaque scénario décrit dans le Tableau VI.8. Quelle que soit la durée de prélèvement, les niveaux dans SCOLA sont globalement plus faibles que dans COLCHIC [0,44 - 0,45] en 2007 pour les mesures de courte durée (1/9 agents chimiques) et les mesures de longue durée (2/15 agents chimiques). Cette différence a tendance à s'atténuer puisqu'en 2015, le constat s'inverse pour les mesures de 30 minutes avec un niveau général plus important dans SCOLA (1,25) pour 6/9 agents chimiques. Pour les mesures sur 240 minutes, les concentrations dans SCOLA restent néanmoins plus faibles que dans COLCHIC pour 10/15 agents chimiques, mais avec un ratio qui s'est réduit (0,85).

Tableau VI.8 Ratios SCOLA/COLCHIC pour les 4 scénarios et de leur intervalle de confiance à 95 % (IC 95 %) par Méta analyse

Scénario	Ratio S/C (IC 95 %)
Situation moyenne en 2007 pour une mesure de :	
30 minutes	0,44 (0,29;0,66)
240 minutes	0,45 (0,32;0,63)
Situation moyenne en 2015 pour une mesure de :	
30 minutes	1,25 (0,99;1,58)
240 minutes	0,85 (0,66;1,10)

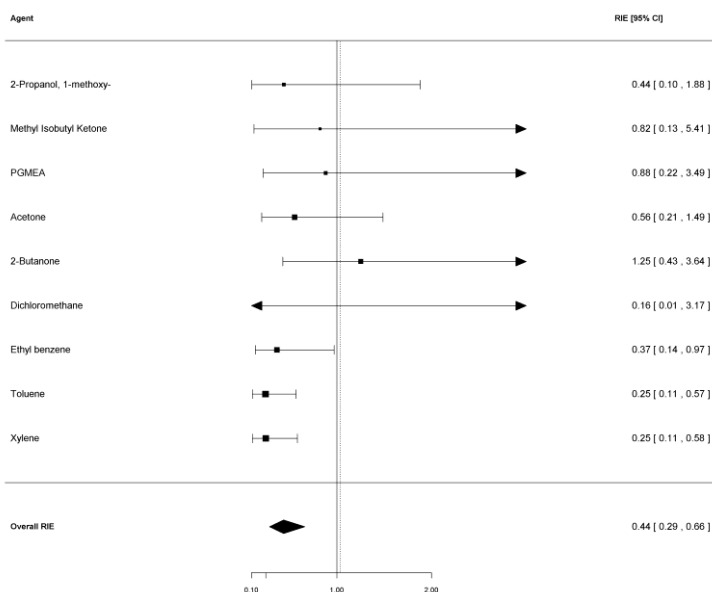
VI.3.3.3. Représentations graphiques des résultats des modèles

Pour chacune des trois analyses de modélisation, les RIE des agents chimiques reliés à l'estimation globale sont présentés graphiquement à l'aide de Forest plots (Figure VI.2). Cette représentation permet de comparer, pour chaque agent chimique et à travers le RIE, une catégorie étudiée et celle définie comme référence. Les carrés noirs, dont la taille est proportionnelle à leurs poids statistiques, représentent le RIE estimé pour une substance. Les lignes horizontales dessinent son IC 95 %. Pour finir, le losange noir correspond à l'estimation globale du RIE issue de la Méta analyse et de son IC 95 %. L'ensemble de ces graphiques est disponible sur <http://expostats.ca/gautier/annexe2.html>. Par exemple, la variabilité interagents chimiques pour les ratios SCOLA/COLCHIC de 2007 est représentée dans les Figures VI.2, A et VI.2, C et celle pour 2015 dans les Figures VI.2, B et VI.2, D.

Les analyses de sensibilité réalisées avec et sans structure ne mettent pas en avant, de manière générale, de variations notables des Méta coefficients (Annexes VI.2 à VI.4).

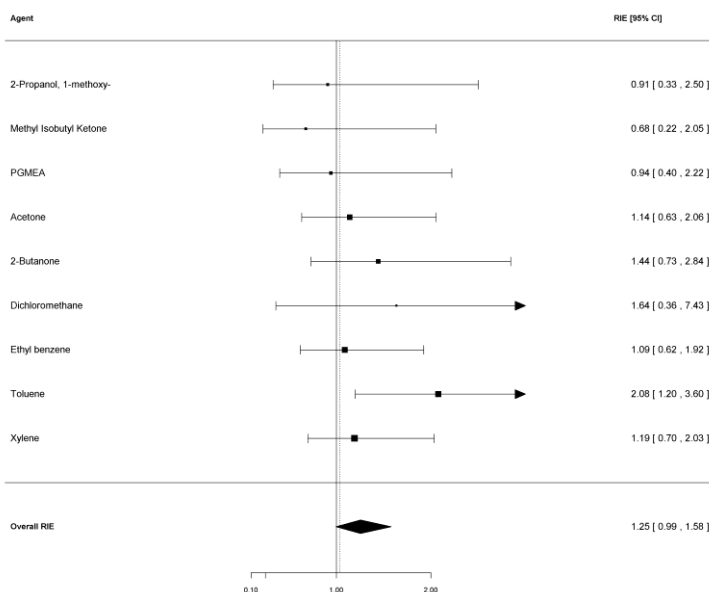
(A)

Effect: diff_SC_2007_30 2007-2015 Q11 TOBIT COLCHIC SCOLA Comparison



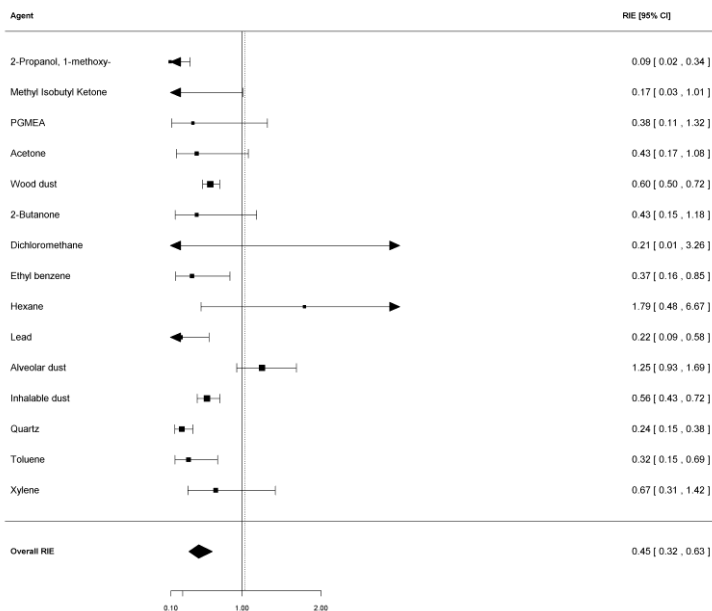
(B)

Effect: diff_SC_2015_30 2007-2015 Q11 TOBIT COLCHIC SCOLA Comparison



(C)

Effect: diff_SC_2007_240 2007-2015 Q11 TOBIT COLCHIC SCOLA Comparison



(D)

Effect: diff_SC_2015_240 2007-2015 Q11 TOBIT COLCHIC SCOLA Comparison

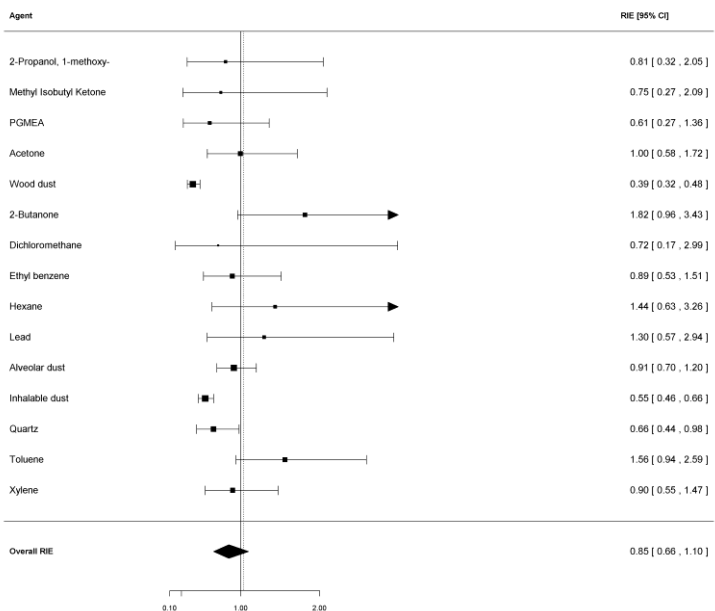


Figure VI.2 : Ratios SCOLA/COLCHIC issus de la Méta analyse par agent chimique pour les 4 scénarios : A- 2007 et durée du prélèvement de 30 minutes, B- 2015 et durée du prélèvement de 30 minutes, C- 2007 et durée du prélèvement de 240 minutes et D- 2015 et durée du prélèvement de 240 minutes

VI.4. DISCUSSION

VI.4.1. MODÉLISATION STATISTIQUE DES DONNÉES D'EXPOSITION

Ce travail représente le premier effort pour étudier globalement l'association entre les expositions et les variables enregistrées dans COLCHIC, SCOLA et la BDEP commune. Les analyses, effectuées pour les mesures individuelles sur la période 2007-2015, ont été stratifiées pour tous les agents chimiques. Une image générale des effets de chaque déterminant d'exposition pour chaque jeu de données a été réalisée grâce à la Méta analyse (Tableaux VI.5, VI.6 et VI.7, Annexes VI.2 et VI.3). Une vue d'ensemble, à travers tous les agents chimiques, est rendue possible par les Forest plots (<http://expostats.ca/gautier/annexe2.html>). Les pourcentages de la variabilité totale, expliqués par les effets fixes pour les analyses séparées de COLCHIC et SCOLA, sont équivalents à ceux obtenus pour leur comparaison [13 % - 56 %]. Nous aurions pu nous attendre à des résultats plus faibles pour la comparaison par rapport aux analyses séparées, puisque seuls les déterminants communs étaient inclus, mais cette tendance ne s'est pas révélée. Ces valeurs sont comparables avec celles de la littérature pour des études similaires [Burstyn et Teschke, 1999 ; Lavoué et al., 2010 ; Sarazin et al., 2016], respectivement [20 % - 70 %], [29 % - 57 %] et [16 % - 24 %].

VI.4.1.1. Analyse séparée de COLCHIC et SCOLA

A travers cette analyse de COLCHIC sur 2007-2015, nous pouvons confirmer d'une manière générale une similitude dans les tendances observées au chapitre V portant sur des périodes plus vastes. Pour éviter les redites, nous nous concentrerons dans cette section sur les résultats de SCOLA en les comparant avec ceux de COLCHIC.

La durée de prélèvement est un prédicteur prépondérant des niveaux mesurés dans COLCHIC et SCOLA. Une augmentation de 30 minutes de la durée de prélèvement a pour conséquence une diminution de la concentration de 34 % et 37 % pour un échantillonnage de courte durée et de 5 % et 4 % pour une mesure de longue durée respectivement pour COLCHIC et SCOLA. Cette tendance, déjà observée lors de l'analyse de COLCHIC sur 1987-

2015 et 2002-2015, se confirme et s'étend dans les mêmes proportions à SCOLA. A travers tous les agents chimiques de SCOLA, les concentrations mesurées sur de longues durées sont plus faibles que celles mesurées sur de courtes durées (facteur 0,39). Ce constat, commun à COLCHIC et SCOLA, est cohérent par rapport à l'attendu puisque d'une manière générale, un prélèvement de courte durée est mis en œuvre pour mesurer une exposition aiguë à un agent chimique alors qu'un prélèvement sur la durée d'un poste de travail a pour objectif d'être représentatif d'une exposition moyenne journalière.

Alors que les niveaux d'exposition baissent annuellement dans COLCHIC [4 % - 7 %] en fonction des périodes, une tendance à la hausse est observée dans SCOLA pour les mesures « Courte durée » (5 %) et une stabilité des concentrations est constatée pour les prélèvements « Longue durée ». Ce constat, globalement différent de celui décrit dans la littérature [Peters et al., 2011 ; Sauvé et al., 2012 ; Sarazin et al., 2016], peut être lié à un biais de sélection : les niveaux enregistrés dans les entreprises échantillonnées ne sont pas représentatifs de ceux de la population générale. La démarche du contrôle réglementaire étant à l'initiative de l'industriel, on peut supposer que les prélèvements effectués dans ce contexte et enregistrés dans SCOLA portaient, les premières années, sur des entreprises avec des niveaux plus faibles.

L'équipement de protection individuelle est fortement associé aux expositions. En effet, lorsqu'il est adapté, le RIE global et son IC 95 % calculés par la Méta analyse n'incluent pas la référence (1,00). Cette importante association entre ce prédicteur et l'exposition, uniforme pour une majorité de substances de COLCHIC et SCOLA, insinue un biais sous-jacent dans les données.

Bien qu'aucune association entre la tranche d'effectif et les concentrations dans COLCHIC sur 2007-2015 et sur 2002-2011 ne soit montrée comme lors des travaux de Clerc et al. [2014], un effet important de cette variable, à travers une large majorité de substances, est remarqué dans SCOLA et corrobore certaines observations décrites dans la littérature [Gomez, 1997 ; Melville et Lippmann, 2001 ; Middendorf, 2004 ; Henn et al., 2011 ; Sarazin et al., 2016]. Cette association est d'autant plus marquée dans SCOLA qu'aucun RIE [IC95 %] de cette variable ne contient la référence (1,00).

Il en est de même pour la ventilation générale qui n'est pas associée à l'exposition dans

nos analyses de COLCHIC comme Clerc et al. [2014] l'avaient déjà constaté. Dans SCOLA, la tendance est différente puisque nous observons une association entre cette variable et l'exposition. En présence d'un captage localisé seul, c'est-à-dire sans ventilation générale, les niveaux mesurés sont plus importants [19 % - 20 %].

Aucune association n'est relevée entre la « Région » et l'exposition dans SCOLA comme cela avait été le cas dans COLCHIC.

Pour finir, nous noterons que la fréquence d'exposition, codée exclusivement dans COLCHIC, est associée à l'exposition sur la période 2007-2015 alors qu'elle ne l'était pas sur 2002-2015. Ainsi, lorsque l'exposition est permanente ou fréquente, les niveaux mesurés dans COLCHIC sont plus élevés que lorsqu'elle est occasionnelle. Cette tendance, vérifiée pour 28/33 substances, montre une importante association entre cet effet et les niveaux d'exposition à travers un grand nombre d'agents chimiques. Ce constat n'avait pas été relevé par Clerc et al [2014] lors de leurs travaux sur les modèles statistiques pour estimer les expositions professionnelles aux agents chimiques à partir des données de COLCHIC sur 2002-2011. En effet, lors de leurs analyses, cette variable avait été classée comme étant la moins prédictive parmi les 13 testées. Nos résultats démontrent que, sur une période plus récente, cette variable est associée aux niveaux d'exposition.

VI.4.1.2. Analyse comparative de COLCHIC et SCOLA

Globalement, pour une majorité de variables communes à COLCHIC et SCOLA, les tendances observées dans les analyses séparées sont identiques à celles issues de la comparaison des 2 BDEP. Mais, ce constat n'est pas vrai pour les variables « Ventilation globale » et « Tranche d'effectif » dont l'association avec les niveaux mesurés est moins cohérente dans l'analyse de comparaison que dans celle de SCOLA. Pour compléter, l'effet de la variable « Source » a montré une différence systématique globale entre les 2 BDEP, avec une différence moyenne de 31 % (COLCHIC > SCOLA) (Annexe VI.1). Cette tendance est relativement uniforme pour 13/15 agents chimiques communs aux 2 BDEP.

Dans le but de comparer les niveaux d'exposition de COLCHIC et SCOLA, les différentes situations décrites à partir des variables influentes moyennes de type catégorie et des

variables influentes de type continue « Année » et « Durée de prélèvement » ont permis deux constats. Premièrement, de confirmer une tendance annuelle à la baisse des concentrations de COLCHIC, quelle que soit la durée de prélèvement et une tendance à la hausse pour les expositions de SCOLA sur les mesures de courte durée, celles de longues durées étant stables dans le temps. Deuxièmement, de corroborer à travers ces scénarios les conclusions de Clerc et Vincent, [2014] qui portent sur la relation entre la durée du prélèvement et le niveau mesuré : une augmentation de la durée de prélèvement est associée à une baisse de la concentration. En 2007, les concentrations enregistrées dans SCOLA sont de l'ordre de 0,45 fois plus faibles que celles de COLCHIC. La tendance à la baisse des niveaux de COLCHIC, conjuguée à une tendance à la hausse ou stagnante des concentrations de SCOLA, engendre un rapprochement implicite des niveaux enregistrés dans chacune de ces bases. En effet, la comparaison des données en 2015 confirme que sur une période récente les niveaux enregistrés dans COLCHIC et SCOLA sont comparables (Annexe VI.5).

Les implications des associations mises en évidence sur l'utilisation de COLCHIC, SCOLA et de la BDEP commune ainsi que les orientations que dégagent nos résultats pour améliorer leur interprétation sont discutées plus en détail dans le chapitre VII.

VI.4.2. LIMITES DE L'ANALYSE

Les limites exposées au chapitre V s'appliquent à cette analyse.

Le démarrage de la réglementation en 2010 a donné un élan pour l'enregistrement des données dans SCOLA. Son caractère récent en limite le nombre d'agents inclus dans la BDEP commune. Lors de futurs travaux, cette restriction devrait néanmoins tendre à disparaître.

L'analyse globale de la base commune ne permettait pas de prendre en compte les interactions avec le secteur d'activité alors que des différences sont plausibles entre COLCHIC et SCOLA. Dans le cadre d'analyse spécifique à un agent chimique donné, il sera important de tenir compte de cet aspect.

VI.5. CONCLUSION

Ce travail a permis d'identifier 3 variables « Durée de prélèvement », « Equipement de protection individuelle » et « Année » avec une association forte et homogène à travers un grand nombre d'agents chimiques, deux autres « Type de procédé » et « Fréquence d'exposition » spécifiques à COLCHIC et enfin deux dernières « Ventilation générale » et « Tranche d'effectif » propres à SCOLA.

Il confirme et étend à SCOLA certaines observations du chapitre V : plus la durée d'un prélèvement est courte plus l'exposition est forte et un travailleur équipé d'une protection individuelle adaptée sera plus fortement exposé que s'il n'est pas équipé.

Néanmoins certaines conclusions divergent : alors que les concentrations mesurées décroissent annuellement dans COLCHIC, une tendance à la hausse et à la stabilité est enregistrée dans SCOLA respectivement pour les prélèvements de courte et de longue durée. Cette augmentation constatée dans SCOLA est vraisemblablement liée à un biais de sélection puisque ce sont les entreprises qui sont à l'origine de la démarche du contrôle.

Ce contexte est néanmoins favorable à un rapprochement des expositions entre les deux BDEP puisque les niveaux historiques de COLCHIC, globalement deux fois plus élevés que ceux de SCOLA, sont comparables ces dernières années. L'agrégation de leurs données pour les agents chimiques concernés permettra de renforcer la connaissance des niveaux *a priori*.

CHAPITRE VII. DISCUSSION GÉNÉRALE

La situation atypique, sur le plan international, de posséder deux bases de données d'exposition professionnelle à des produits chimiques et biologiques est une particularité française. COLCHIC et SCOLA, deux sources importantes d'information sur les niveaux d'exposition *a priori*, sont ainsi disponibles en France. Leurs différents objectifs, qui orientent les stratégies de prélèvements déployées en entreprise, en font leur originalité : COLCHIC, alimentée par les huit laboratoires interrégionaux de chimie des CARSAT/CRAM et l'INRS, enregistre des données mesurées dans un cadre de prévention alors que SCOLA est exclusivement portée sur le volet réglementaire à travers les concentrations saisies par les organismes accrédités dans le domaine de l'air des lieux de travail. Malgré ces différences, le terrain d'investigation est commun et constitue la spécificité de la situation. Avec une architecture technique commune, les informations servant à décrire les situations de travail sont codées dans COLCHIC et SCOLA à l'aide de tables de références en grande majorité communes (Chapitre III). Dans ce contexte favorable, cette recherche visait dans un premier temps à identifier les variables associées systématiquement aux niveaux d'exposition mesurés et les biais potentiels dans chacune des deux BDEP. Dans un deuxième temps, elle cherchait à étudier les différences et les similitudes entre COLCHIC et SCOLA en tenant compte des associations préalablement mises en évidence, pour finalement apporter des éléments de réponse à la question « Dans quelle mesure les données d'exposition à des produits chimiques, enregistrées dans COLCHIC et SCOLA reflètent-elles les expositions professionnelles de façon historique en France ? ».

Tableau VII.1 : Synthèse des principaux résultats obtenus lors des analyses des bases de données COLCHIC et SCOLA

Chapitre IV : Analyse descriptive et comparaison de deux bases de données d'exposition professionnelle françaises COLCHIC (C) et SCOLA (S)

Méthodes d'analyse

- Bases de données COLCHIC (C) et SCOLA (S)
- Comparaison descriptive (variables, lieux de travail, agents chimiques et niveaux d'exposition)
- Périodes 1987-2012 (C) et 2007-2012 (S)

Principaux résultats

- COLCHIC 841 682 résultats, 670 agents chimiques (AC)
- SCOLA 152 486 résultats, 70 AC
- Individuel/Ambiant : 70 %/30 % (C) – 97 %/3 % (S)
- LoQ/Déecté : 31 %/69 % (C) – 88 %/12 % (S)
- Durée prélèvement individuel médiane 2007-2012 (minutes) : 28 et 15 courte durée - 215 et 377 longue durée respectivement (C) et (S)
- En commun, NAF 69 %, Métier 70 %, Tâche 65 %, 17 agents chimiques (> 500 mesures)
- Ratio Conc. médiane (C)/(S) 3,45 [1,03-14,3] pour 17 AC communs

Chapitre V : Analyse par modélisation de la base de données COLCHIC

Méthodes d'analyse

- Périodes 1987-2015 (87-15) et 2002-2015 (02-15)
- Type prélèvement : Individuel et Ambiant
- 86 (87-15) et 50 (02-15) substances (>1 000 mesures)
- 15 (87-15) et 18 (02-15) variables analysées
- Modèle TOBIT (% LoQ<60)
- ~780 000 (87-15) et ~335 000 (02-15) mesures
- Synthèse résultats à travers les substances par Méta analyse
- Variabilité intersubstances : Forest Plots

Principaux résultats

- Mesures longue durée < Mesures courte durée (facteur [0,46-0,64] selon période et type prélèvement)
- Tendance annuelle : réduction [4 %-10 %] par an selon la période et le type de prélèvement
- Niveaux EPI adapté > Niveaux EPI absent ou non adapté (facteur [1,39-1,86] selon période et type prélèvement)
- Niveaux autres types procédé < Niveaux procédé ouvert, réduction [4 %-28 %] selon type prélèvement

Chapitre VI : Modélisation comparative des bases de données COLCHIC et SCOLA

Méthodes d'analyse

- Bases de données COLCHIC (C), SCOLA (S) et jeu de données commun COLCHIC + SCOLA (CS)
- Période 2007-2015
- Type prélèvement : Individuel
- 39 (C), 35 (S) et 15 (CS) substances (>1 000 mesures)
- 17 (C), 13 (S) et 14 (CS) variables analysées
- Modèle TOBIT (% LoQ<60)
- ~170 000 (C) ~208 000 (S) et ~240 000 (CS) mesures
- 4 scénarios de prédiction pour le ratio (S/C) : années 2007 et 2015, durée prélèvement 30 et 240 minutes
- Synthèse résultats à travers substances par Méta analyse
- Variabilité intersubstances : Forest Plots

Principaux résultats

- Mesures longue durée < Mesures courte durée (facteur ~0,5)
- Tendance annuelle : réduction 7 % (C), augmentation 5 % courte durée - stable longue durée (S). Tendances identiques (CS)
- Niveaux EPI adapté > Niveaux EPI absent ou non adapté (facteur ~1,8 (C), ~2,5 (S)) et tendances identiques (CS)
- Variables « Type de procédé » et « Fréquence d'exposition » influentes (C)
- Variables « Ventilation générale » et « Tranche effectif » influentes (S)
- Avant 2013, Niveaux (C) > (S), après 2013 Niveaux (C) et (S) comparables courte durée, dès 2014 longue durée
- Ratio (S/C) en 2007 0,44 (30 min) et 0,45 (240 min) en 2015 1,25 (30 min) et 0,85 (240 min)

VII.1.MODÉLISATION STATISTIQUE DES DONNÉES D'EXPOSITION

VII.1.1. GÉNÉRALITÉS

Pour identifier les variables associées aux niveaux d'exposition dans COLCHIC et SCOLA, nous avons été amenés dans le cadre de ces travaux de thèse à stratifier les analyses :

- 1- selon trois périodes, 1987-2015 et 2002-2015 pour COLCHIC et 2007-2015 pour COLCHIC, SCOLA et la BDEP commune
- 2- par type de prélèvement « individuel » pour tous les jeux de données et « ambiant » uniquement pour les deux grandes périodes de COLCHIC
- 3- pour tous les agents chimiques présents dans chaque jeu de données et par famille de solvants et de métaux lorsque le nombre d'agents chimiques était suffisant.

Après avoir identifié les variables d'intérêt présentes dans COLCHIC et SCOLA (Chapitre IV), nous avons étudié les associations entre les variables d'intérêt et les niveaux mesurés pour chaque jeu de données à partir de modèles statistiques de régression censurés [Tobin, 1958 ; Lubin et al., 2004 ; Lavoué et al., 2010 ; Sauvé et al., 2012] (Chapitres V et VI). Compte tenu de la quantité très importante de résultats générés, nous nous sommes inspirés des travaux de Sarazin et al. [2015] pour présenter une image générale des effets de chaque déterminant d'exposition pour chaque strate à travers l'ensemble des agents disponibles. Ces chercheurs sont les pionniers dans le domaine de l'expologie pour avoir employé la Méta analyse, largement utilisée dans le domaine de l'épidémiologie, pour synthétiser leurs résultats. Nous avons ensuite renforcé nos résultats de Méta analyse en étudiant la variabilité inter agents chimiques à l'aide de Forest plots disponibles directement à partir de ces adresses : <http://expostats.ca/gautier/annexe1.html> et <http://expostats.ca/gautier/annexe2.html>.

Une synthèse des principaux résultats obtenus classés par chapitre est présentée dans le Tableau VII.1. Ces résultats sont discutés dans les sections suivantes.

VII.1.2. COLCHIC 1987-2015, 2002-2015 ET 2007-2015

Afin de tenir compte d'une part, des spécificités liées à la codification de nouvelles variables en 2002 et d'autre part, de la période commune aux deux BDEP, la stratification

des mesures de COLCHIC a généré 5 jeux de données (1987-2015 et 2002-2015 individuel et ambiant puis 2007-2015 individuel pour la comparaison avec SCOLA). Pour la période 2007-2015, l'objectif de l'analyse étant de comparer COLCHIC à SCOLA, seules les données individuelles ont été exploitées. A travers ces analyses, les pourcentages de variabilité totale expliqués par les effets fixes [13 %-69 %] sont comparables avec ceux relevés dans la littérature pour des travaux similaires [Burstyn et Teschke, 1999 ; Lavoué et al., 2005 et 2010 ; Sarazin et al., 2016] et confortent la qualité de notre étude.

Lors de nos travaux, nous avons pu observer que l'association d'une variable avec l'exposition était dans la majorité des cas constante, quelle que soit la période investiguée. Dans ce contexte, les variables présentant des effets similaires renforcent les conclusions sur leur association avec l'exposition.

Le ratio entre un prélèvement de longue durée et un prélèvement de courte durée est similaire pour les 5 jeux de données [0,46-0,64]. Ce constat conforte ceux établis par Raaschou-Nielsen [2002], Lavoué et al., [2005] et Hon et al., [2016]. Ainsi, plus un prélèvement est long plus la concentration mesurée est faible. Ce propos, déjà rapporté pour le toluène par Clerc et Vincent [2014], est maintenant vérifié à travers tous les agents chimiques de COLCHIC. Ce phénomène laisse présumer un biais d'échantillonnage qui peut s'expliquer à travers ces deux situations : 1- prélèvement court pour mesurer l'exposition à une tâche particulière et ainsi connaître l'exposition aiguë de l'opérateur 2- prélèvement long pour connaître l'exposition journalière du travailleur et ainsi estimer son exposition chronique. Annuellement, l'exposition diminue [4 %-10 %] quels que soient la période considérée, les types de prélèvement et d'échantillonnage. Cette décroissance, déjà observée dans la littérature lors de travaux équivalents [Gomez et al., 1997 ; Symanski et al., 1998 et 2001 ; Okun et al., 2004 ; Lavoué et al., 2005], est vraisemblablement le fruit des politiques de prévention mises en œuvre en général et plus particulièrement en France dans notre cas. L'équipement de protection individuelle est fortement associé à l'exposition à travers une grande majorité d'agents chimiques. Le port d'un équipement adapté est associé avec un niveau d'exposition plus élevé que lorsqu'il est absent [1,39-1,86]. Cette variable, fortement associée à la concentration, est un important prédicteur pour une large majorité de substances à travers toutes les périodes. L'exposition tend à être élevée dans les endroits où un salarié est amené à

porter une protection. Enfin, la variable « Type de procédé » est fortement associée avec l'exposition. Lorsqu'un procédé est ouvert, les concentrations mesurées sont toujours plus importantes que dans les autres situations (semi-ouvert ou fermé).

Dans d'autres cas, les variables ne sont pas associées aux expositions. Par exemple, pour les variables « Région », « Motif de la demande », « Origine de la demande », « Ventilation générale » ou encore « Tranche d'effectif », aucune association globale avec les niveaux de concentration est constatée. Pour certaines d'entre elles, ces conclusions confortent celles issues de la littérature [Froines et al., 1986 et 1990 ; Gomez, 1997 ; Melville et Lippmann, 2001 ; Middebdorf, 2004 ; Lavoué et al., 2005, 2008 et 2013 ; Henne et al., 2011 ; Sarazin et al., 2016]. Cependant, nous noterons que nos conclusions sont globales alors que dans le cadre de ces travaux, les chercheurs ne se sont intéressés qu'à des agents particuliers. Ainsi, si nous nous focalisons sur un effet spécifique, le « Motif de la demande » pour un agent chimique particulier comme le quartz, nous observerions une forte association entre ce descripteur et l'exposition pour les périodes 1987-2015 et 2002-2015 alors que globalement, à travers un grand nombre d'agents chimiques aucune association n'est observée.

Enfin, dans une situation, l'association entre la variable et le niveau a évolué d'un jeu de données à un autre. En effet, pour la « Fréquence d'exposition » sur 2002-2015, aucune association n'est observée comme Clerc et al. [2014] l'avaient aussi constaté alors que sur 2007-2015 l'association est homogène pour 28/33 agents chimiques.

Pour résumer, ces analyses ont permis d'estimer l'effet sur les niveaux mesurés de 8 variables descriptives pour la période 1987-2015 et de 10 pour les périodes 2002-2015 et 2007-2015. Quatre variables ont une forte association sur les niveaux enregistrés à travers un grand nombre d'agents chimiques : « Durée de prélèvement », « Année », « Equipement de protection individuelle » et « Type de procédé », quels que soient la période et le type de prélèvement. La variable « Fréquence d'exposition » est associée aux expositions uniquement sur la période la plus récente. Deux autres variables « Origine de la demande » et « Ventilation générale » ont une certaine influence sur les niveaux d'exposition, mais leur portrait est moins cohérent à travers toutes les données de COLCHIC. Enfin, les trois dernières variables « Région », « Motif de la demande » et

« Tranche d'effectif » ont, d'une manière générale, une faible association sur les niveaux mesurés. Pour ces 6 dernières variables, il est plus difficile de donner une interprétation globale.

Ces résultats montrent donc la nécessité de prendre en compte les variables identifiées lors de l'interprétation de COLCHIC. Pour les variables sans tendance globale, elles apparaissent moins essentielles, mais pourraient être importantes pour des agents particuliers.

VII.1.3. SCOLA 2007-2015

Les analyses par modélisation de SCOLA ont été effectuées pour tous les agents chimiques sur un seul jeu de données portant sur les mesures individuelles pour la période 2007-2015. Les pourcentages de la variabilité totale expliqués par les effets fixes [15 %-56 %] sont comparables avec ceux de COLCHIC et ceux des études similaires identifiées dans la littérature [Burstun et Teschke, 1999 ; Lavoué et al., 2010 ; Sarazin et al., 2016].

Les tendances observées pour les effets de SCOLA sont généralement proches de celles obtenues pour COLCHIC. Ces similitudes globales entre ces deux BDEP, qui possèdent des stratégies de sélection des entreprises et d'échantillonnage très différentes, renforcent la fiabilité, pour ces variables, dans les estimés des niveaux enregistrés dans chacune d'entre elles. Ainsi, on retrouve l'équipement de protection individuelle qui est aussi fortement associé à l'exposition dans SCOLA et de manière homogène à travers tous les agents chimiques. De même, l'effet lié à la durée de prélèvement présente aussi des conclusions de tendances à la baisse et dans de mêmes proportions dans SCOLA [37 % courte durée, 4 % longue durée] que dans COLCHIC [34 % courte durée, 5 % longue durée] sur 2007-2015. Enfin, pour la variable « Région » l'absence d'effet remarquée dans COLCHIC est confirmée dans les analyses de SCOLA.

Certaines différences systématiques sont néanmoins apparues entre COLCHIC et SCOLA et suggèrent un effet de stratégie globale d'échantillonnage. Une première différence s'illustre à travers l'augmentation annuelle des concentrations dans SCOLA (5 %) pour les

mesures de courte durée et sur une stagnation pour celles de longue durée alors que dans COLCHIC une baisse de 7 % des concentrations est observée pour la même période. Cette tendance à la hausse dans SCOLA peut s'expliquer d'une part, parce que les campagnes de mesures réglementaires sont réalisées à la demande de l'entreprise. Cette approche laisse présumer que SCOLA a commencé à être renseignée par les mesures des campagnes réglementaires pour les entreprises confiantes dans leur politique hygiène et sécurité puis, au cours des années, avec une participation plus large d'entreprises devant répondre à leurs obligations réglementaires supposant donc des niveaux plus faibles les premières années. D'autre part, les données enregistrées dans SCOLA, avant la mise en application du décret 2009-1570 du 15 décembre 2009, étaient saisies sur la base du volontariat des organismes accrédités. Ces constats induisent implicitement un biais de sélection des données. Une deuxième différence met en évidence une association forte entre la variable « Tranche d'effectif » et les expositions dans SCOLA pour une large majorité d'agents chimiques. Les niveaux enregistrés sont inversement proportionnels à la taille de l'entreprise. Ainsi, une entreprise comptant 1 à 19 salariés présentera des concentrations plus importantes que celle de plus de 200 salariés (35 %). Cette association corrobore les conclusions de Gomez [1997], Melville et Lippmann [2001], Middendorf [2004] Henn et al., [2011] et Sarazin et al., [2016] bien qu'une absence d'effet soit notée pour COLCHIC. Cette différence d'association entre COLCHIC et SCOLA est vraisemblablement liée aux objectifs de chacune des deux BDEP (les préventeurs ciblent préférentiellement les entreprises de petites tailles ; la réglementation est mise en œuvre plus rapidement dans les grandes entreprises qui ont plus de moyens) et présume d'un effet de stratégie globale d'échantillonnage. Enfin, une troisième différence porte sur la ventilation générale qui est associée aux expositions dans SCOLA pour 13/20 agents chimiques alors qu'elle ne l'est globalement jamais dans COLCHIC. En présence d'un captage à la source seul, les niveaux enregistrés dans SCOLA sont les plus importants (20 %) alors que dans COLCHIC ce phénomène est observé en présence d'une ventilation générale seule (8 % comparé à la référence, captage localisé sans ventilation générale). Deux tendances peuvent s'opposer : si nous calquons le raisonnement de l'équipement de protection individuelle, la présence d'une ventilation est indicatrice de milieux exposés. A l'inverse, en présence d'une ventilation et/ou d'un captage à la source, les

niveaux devraient être faibles. Là encore, la sélection des lieux échantillonnés peut faire jouer différemment ces tendances.

Ces analyses par modélisation des données enregistrées dans SCOLA ont permis d'évaluer l'effet de 6 variables sur les niveaux d'exposition. Il en ressort 5 variables fortement associées aux niveaux enregistrés dans SCOLA pour une large gamme d'agents chimiques « Durée de prélèvement », « Equipement de protection individuelle », « Année », « Tranche d'effectif » et « Ventilation générale ». Une seule variable « Région » n'a pas montré d'association globale avec les niveaux et rend son interprétation plus difficile. Cette variable semble donc moins essentielle dans notre situation, mais pourrait s'avérer être importante lors d'analyses avec des agents particuliers.

VII.1.4. COMPARAISON DES DONNÉES CONJOINTES DE COLCHIC ET SCOLA 2007-2015

Le principal objectif de cette section consiste à comparer les niveaux d'exposition moyens entre COLCHIC et SCOLA.

Lors de l'analyse du modèle simple, sans prise en compte des interactions, la variable « Source » a montré une forte association avec les expositions à travers une large gamme d'agents chimiques (12/15). Il en ressort des niveaux enregistrés dans SCOLA globalement plus faibles que dans COLCHIC (31 %) sur la période 2007-2015 (Annexe VI.4) et confirme la tendance déjà observée dans le chapitre III.

Les tendances observées pour le modèle avec interaction reflètent bien celles que l'on a vues dans les analyses séparées pour les variables « Année », « Durée du prélèvement », « Equipement de protection individuelle » et « Région ». Concernant les variables « Ventilation générale » et « Tranche d'effectif », malgré une forte association observée entre ces deux variables et l'exposition dans SCOLA seule, les résultats du modèle avec interaction diffèrent de ceux des analyses séparées. Cette différence pourrait s'expliquer par le fait que, d'une part, l'analyse comparative est restreinte aux agents chimiques communs et, d'autre part, nous n'avons pas pu modéliser d'interactions avec l'activité, ce qui pourrait avoir causé une confusion.

Les 4 scénarios définis sur 2007 et 2015 pour des durées de prélèvement de 30 et 240 minutes ont permis d'observer, sur les premières années de la comparaison, des différences systématiques entre les niveaux enregistrés dans COLCHIC et SCOLA (2007, 0,44 et 0,45 respectivement 30 et 240 minutes). Puis, nous avons pu observer que cette tendance s'atténue au fil des années jusqu'à arriver à des niveaux similaires observés entre COLCHIC et SCOLA pour les prélèvements de courte durée à partir de l'année 2013 et pour ceux de longue durée à partir de l'année 2014 (2015, 1,25 et 0,85 respectivement 30 et 240 minutes) (Annexe VI.5). Si cette tendance se confirme, la différence entre COLCHIC et SCOLA devrait se stabiliser. En plus de ces similitudes dans les niveaux mesurés, ces observations laissent présager la possibilité d'utiliser de manière combinée les données issues de COLCHIC et SCOLA pour les agents chimiques communs dans la mesure où cette tendance perdure. Elles pourront ainsi servir de source d'information sur l'exposition dès lors que les mesures portant sur les premières années de leur période commune auront été exclues.

VII.2. UTILISATION DES DONNÉES DE COLCHIC ET SCOLA POUR ÉVALUER LES NIVEAUX DE LA POPULATION GÉNÉRALE

Les travaux de modélisation décrits précédemment ont entre autres permis d'identifier, à travers les différents jeux de données, les variables associées systématiquement ou partiellement aux niveaux d'exposition. Nous avons ainsi pu relever que 3 prédicteurs « Année », « Durée de prélèvement » et « Equipement de protection individuelle » avaient une tendance claire et globale à travers une large majorité d'agents chimiques pour COLCHIC et SCOLA. D'autres prédicteurs sont spécifiques à une base de données, soit parce qu'ils ne sont pas codés dans l'autre base et que leur association n'a pas pu être évaluée, par exemple le « Type de procédé » et la « Fréquence d'exposition » pour COLCHIC, soit parce que la différence de stratégie globale d'échantillonnage entre COLCHIC et SCOLA engendre une différence d'association entre le prédicteur et l'exposition dans chacune des bases de données (« Ventilation générale » et « Tranche d'effectif » dans SCOLA). Pour les autres variables non citées précédemment, aucune association globale à travers les agents chimiques n'a été observée.

Pour qu'une BDEP soit représentative, la distribution des mesures enregistrées doit refléter celle définie dans la population générale. Prenons comme exemple le secteur de l'industrie manufacturière en France. En 2014, environ 70 % des entreprises de cette branche d'activité sont des micro-entreprises (1 à 9 salariés). Les 30 % restant, que nous qualifierons de grandes entreprises dans cet exemple, ont un effectif supérieur à 10 salariés (Source, www.insee.fr). Nous avons identifié lors de nos analyses que la « Tranche d'effectif » est fortement associée aux niveaux enregistrés dans SCOLA. S'il était démontré que les concentrations sont systématiquement plus fortes dans les micro-entreprises de ce secteur, il pourrait être nécessaire d'ajuster les prédictions issues des données de SCOLA. Deux orientations sont possibles : si la répartition des mesures par tranches d'effectif dans SCOLA pour ce secteur d'activité est identique à celle de la population générale du secteur de l'industrie manufacturière, alors il ne sera pas nécessaire d'ajuster les prédictions. Par contre, étant donné que la répartition dans SCOLA est différente, environ 10 % de micro-entreprises et 90 % de grandes entreprises, un ajustement des prédictions faites à partir du contenu de SCOLA s'avère nécessaire. En effet, si cette différence de distribution n'est pas prise en compte, une sous-estimation des expositions réelles serait causée pour ce secteur. Connaissant la distribution réelle des tranches d'effectif, il est alors possible de pondérer les estimés d'exposition issus de SCOLA en fonction de la distribution donnée par l'INSEE.

Pour d'autres variables, il est plus difficile de connaître leur distribution à l'intérieur de la population générale. Prenons comme exemple le prédicteur « Equipement de protection individuelle ». Ce dernier est très fortement associé aux niveaux mesurés que ce soit dans COLCHIC ou SCOLA. Nous avons observé que les concentrations sont bien plus importantes lorsque le salarié est équipé d'une protection adaptée que lorsqu'il n'en porte pas, ce qui suggérerait d'ailleurs un biais de sélection (chaque situation conduisant à une exposition particulière). Là encore, il paraîtrait nécessaire d'ajuster les prédictions faites à partir du jeu de données. Mais dans ce cas, il est difficile de savoir dans quelle proportion chacune de ces situations reflète celle de la population générale. Pour effectuer une correction des estimés, il conviendra alors de s'appuyer sur l'expérience et/ou l'expertise soit, pour porter un jugement permettant de définir la situation la plus représentative de la population générale soit, pour définir arbitrairement un poids à

chacune de ces catégories. Par exemple, pour le secteur du « Commerce de détail de carburants en magasin spécialisé », nous pourrions définir que la proportion de salariés portant une protection adaptée est faible, lui affecter un poids de 15 % et un poids de 85 % pour l'absence de protection pour la pondération de nos estimés.

Plus concrètement, nous souhaitons dans ces conditions estimer le niveau de chrome VI dans la population générale du secteur de l'industrie manufacturière. Si nous moyennons directement les concentrations enregistrées dans SCOLA pour ce secteur d'activité, l'exposition moyenne serait de $15 \mu\text{g.m}^{-3}$. Hors, connaissant la distribution des entreprises dans la population générale de ce secteur (70 % micro-entreprises et 30 % grandes entreprises), nous pouvons pondérer les concentrations de SCOLA (moyenne de $5 \mu\text{g.m}^{-3}$ dans les micro-entreprises et moyenne de $16 \mu\text{g.m}^{-3}$ dans les grandes entreprises) en tenant compte de la répartition de la population générale. Son exposition moyenne s'estime alors à $8 \mu\text{g.m}^{-3}$. Dans ce contexte, le niveau moyen de SCOLA surestimerait celui de la population générale de ce secteur. Cela s'explique par une proportion importante de grandes entreprises dans SCOLA (90 %) avec des niveaux près de 3 fois supérieurs à ceux des micro-entreprises alors que dans la population générale seulement 30 % de grandes entreprises sont recensées.

VII.3. LIMITES DES ANALYSES

Cette section abordera les limites générales de nos travaux de thèse. Celles spécifiques aux analyses ont été présentées dans les chapitres précédents.

L'ensemble des analyses réalisées au cours de ce travail permettent de dresser un portrait général de l'association des variables d'intérêt avec les niveaux d'exposition, mais ne sont pas spécifiques à un agent chimique particulier. Le choix des structures « Activité économique », « Tendance temporelle et Durée de prélèvement » et « Aléatoire » est issu de compromis permettant de répondre à nos objectifs. Dans ce contexte, l'approche employée pour construire les modèles repose sur les critères de qualité du plus grand nombre de substances. De plus, bien que des différences soient plausibles entre COLCHIC et SCOLA, l'analyse globale de la base commune ne permettait pas de prendre en compte les interactions avec le secteur d'activité. Ce sont pour ces

raisons que dans le cadre de futurs travaux portant sur un agent particulier, un ajustement du modèle par rapport aux objectifs de l'étude sera nécessaire.

Après avoir identifié, lors de l'analyse descriptive, les variables jugées d'intérêt pour COLCHIC et SCOLA, il est apparu pertinent de toutes les conserver dans les modèles statistiques. L'utilisation de modèles complets pour chaque agent peut engendrer des possibilités d'observer des associations dues à la chance. Cependant ces éventuelles associations dues à la chance auraient disparu avec la Méta analyse. Lors de futures analyses, pour établir des portraits pour des agents chimiques particuliers, il faudra envisager des méthodes de sélection de modèle [Lavoué et Droz, 2009].

L'intersection COLCHIC/SCOLA encore limitée principalement à cause de la jeunesse de la base de données SCOLA (105 agents chimiques) conjuguée à l'absence de cadre pour l'enregistrement des données dans SCOLA avant 2010 engendre naturellement un biais de sélection qui devrait tendre à disparaître une fois qu'un régime stable sera établi (Figure VI.1). Cette limite met en avant l'importance de renouveler cet exercice dans le futur, pour vérifier cette hypothèse.

L'absence de « gold standard », c'est-à-dire d'un échantillon aléatoire et représentatif de l'exposition des salariés en France, nous a naturellement orientés vers une approche indirecte qui consiste en l'évaluation des associations avec les variables enregistrées dans COLCHIC et SCOLA et la BDEP commune. Bien que cette approche nous ait permis, par la comparaison des niveaux enregistrés dans chacune des deux BDEP, de mettre en évidence des différences systématiques, elle ne permet pas de conclure sur la représentativité des mesures par rapport à la population générale. Néanmoins, l'identification des variables associées systématiquement aux niveaux devrait permettre de corriger leurs influences et d'améliorer la fiabilité des estimations.

VII.4. ORIGINALITÉ DE CES TRAVAUX DE RECHERCHE

Ce travail porte sur la totalité des données enregistrées dans deux grandes bases de données d'exposition professionnelle françaises à des produits chimiques comptant près de 1,4 million de résultats (68 % COLCHIC et 32 % SCOLA).

Historiquement, certains travaux ciblés sur un agent chimique particulier [Lavoué et al., 2006] ou sur une période plus restreinte [Clerc et al. 2015] avaient permis d'identifier certaines associations spécifiques. L'originalité de ce travail réside, entre autres, dans la mise en œuvre d'une approche commune permettant d'une part, une analyse globale à travers tous les agents chimiques pour différentes périodes englobant la période la plus complète et d'autre part, d'observer la variabilité inter agents chimiques à travers les Forest plots.

COLCHIC renseignée par le réseau de prévention et SCOLA par les organismes accrédités pour les contrôles réglementaires ont chacune un objectif qui leur est propre, mais compilent toutes les deux des mesures d'exposition provenant d'un même tissu industriel. Cette situation atypique est unique sur le plan international et nous a offert l'opportunité de comparer les niveaux enregistrés dans chacune de ces bases.

VII.5. PERSPECTIVES ET RECOMMANDATIONS

Chaque étape conduisant à la mesure puis à son enregistrement dans une BDEP générera potentiellement un biais. Dans ce contexte, fort est de constater que les bases de données ne permettent pas d'établir directement un portrait représentatif de la population générale. Cependant, les résultats apportés par ces travaux de thèse seront précieux pour de futures exploitations pour dresser par exemple, un portrait multi-sectoriel de l'exposition à une substance donnée. Pour des variables identifiées comme étant associées aux expositions pour une majorité de substances dans COLCHIC et SCOLA, une attention particulière devra être portée puisqu'elles présument de tendance sous-jacente dans ces BDEP. Ces prédicteurs pourront aussi être inclus dans des applications prédictives des niveaux pour une substance donnée. Par exemple aujourd'hui, les applications SOLVEX et FIBREX fournissent, à partir d'une requête simple portant sur la période considérée, l'activité économique et la substance d'intérêt, une analyse statistique des niveaux enregistrés dans COLCHIC. En s'appuyant sur nos résultats de recherche, ces applications pourront d'une part, gagner en confiance dans leurs évaluations puisque nous pourrions, pour les agents chimiques communs à COLCHIC et SCOLA, agglomérer leurs données et ainsi renforcer la robustesse de l'information.

D'autre part, les variables identifiées comme associées aux niveaux permettront, moyennant une évolution de ces applications, de prédire des expositions pour des situations qui seraient décrites par les utilisateurs à partir des variables d'intérêt proposées.

Bien que les tendances des niveaux enregistrés dans COLCHIC et SCOLA se soient rapprochées ces dernières années, il est difficile aujourd'hui de garantir que leurs données soient toujours comparables dans le futur. Dans ce contexte, il est fortement recommandé de surveiller leurs évolutions afin de s'assurer que leur exploitation conjointe garde tout son sens.

La variable « Type de procédé » a montré au travers des analyses de COLCHIC une forte association pour un grand nombre d'agents chimiques. Il nous paraît judicieux au regard de ces résultats de recommander que ce prédicteur soit intégré à la codification des prélèvements dans SCOLA. De manière plus générale, plus une BDEP contient d'informations, meilleures seront les futures interprétations.

CHAPITRE VIII. CONCLUSION GÉNÉRALE

Ce travail de thèse a permis, après avoir identifié les prédicteurs dans les analyses séparées de COLCHIC et SCOLA, de comparer ces deux BDEP dont l'origine des mesures est issue d'un même tissu industriel, mais avec chacune leur spécificité en termes d'objectifs (prévention et réglementation). Nous avons mis en évidence des associations cohérentes à travers les agents, qui pourraient causer des biais lors de l'établissement de portraits d'exposition. Elles doivent être prises en compte pour rapprocher le plus possible les estimations des niveaux d'exposition réels en milieu de travail en France. L'amplitude modérée de ces associations et la similarité des niveaux pour les années récentes dans les deux bases de données restent rassurantes à cet égard, malgré l'absence d'étalon or ou « gold star » auquel comparer COLCHIC et SCOLA. La convergence des niveaux de COLCHIC et SCOLA ouvre des perspectives pour de futures exploitations conjointes renforçant ainsi la connaissance *a priori* des expositions pour les substances communes à ces deux bases.

RÉFÉRENCES

- ACGIH-AIHA. 1996. Data Elements for Occupational Exposure Databases: Guidelines and Recommendations for Airborne Hazards and Noise. *Appl Occup Environ Hyg* 11: 1294-1311.
- Bégin D, Gérin M, Adib G. 1995. Development of an Occupational exposure Data Bank on the Territory of a Department of Community Health in Montréal. *Appl Occup Environ Hyg* 10: 355-360.
- Burns D, Beaumont P. 1989. The HSE National Exposure Database (NEDB). *Ann Occup Hyg* 33: 1-14.
- Carton B. 1995. COLCHIC Chemical Exposure Database: Information on Lead and Formaldehyde. *Appl Occup Environ Hyg* 10: 345-350.
- Clerc F, Bertrand N, Vincent R. 2015. TEXAS: a Tool for EXposure ASsessment—Statistical Models for Estimating Occupational Exposure to Chemical Agents. *Ann Occup Hyg* 59: 277-291.
- Clerc F, Eypert Blaison C, Guimon M, Romero-Hariot A, Vincent R. 2011. Campagne de mesures d'exposition aux fibres d'amiante par microscopie électronique à transmission analytique (META): INRS. p 1-164.
- Clerc F, Steinhausen M, Bertrand N, Vincent R, Gabriel S, Van Gelder R. 2015. Comparison of formaldehyde exposure measurements stored in French and German databases. *Gefahrstoffe - Reinhaltung der Luft* 75 (2015) Nr. 4 - April: 119-126.
- Clerc F, Vincent R. 2014. Assessment of Occupational Exposure to Chemicals by Air Sampling for Comparison With Limit Values: The Influence of Sampling Strategy. *Ann Occup Hyg* 58: 437-449.
- Coble J, Lees P, Matanoski G. 2010. Time Trends in Exposure Measurements from OSHA Compliance Inspections of the Pulp and Paper Industry. *Appl Occup Environ Hyg* 16: 263-270.

- Databases A-ATGoOE. 1996. Data Elements for Occupational Exposure Databases: Guidelines and Recommendations for Airborne Hazards and Noise. *Appl Occup Environ Hyg* 11: 1294-1311.
- Driscoll TR, Carey RN, Peters S, Glass DC, Benke G, Reid A, Fritschi L. 2016. The Australian Work Exposures Study: Occupational Exposure to Lead and Lead Compounds. *Ann Occup Hyg* 60: 113-123.
- Févotte J, Dananché B, Delabre L, Ducamp S, Garras L, Houot M, Luce D, Orlowski E, Pilorget C, Lacourt A, Brochard P, Goldberg M, Imbernon E. 2011. Matgene: A Program to Develop Job-Exposure Matrices in the General Population in France. *Ann Occup Hyg*.
- Flanagan ME, Seixas N, Becker P, Takacs B, Camp J. 2006. Silica Exposure on Construction Sites: Results of an Exposure Monitoring Data Compilation Project. *J Occup Environ Hyg* 3: 144-152.
- Froines JR, Wegman DH, Dellenbaugh CA. 1986. An approach to the characterization of silica exposure in U.S. industry. *Am J Ind Med* 10: 345-361.
- Hamm M, Burstyn I. 2011. Estimating Occupational Beryllium Exposure from Compliance Monitoring Data. *Arch Environ Occup Health* 66: 75-86.
- Helsel D. 2012. *Statistics for Censored Environmental Data Using Minitab and R*, 2nd Edition.
- Henn SA, Sussell AL, Li J, Shire JD, Alarcon WA, Tak S. 2011. Characterization of lead in US workplaces using data from OSHA's integrated management information system. *Am J Ind Med* 54: 356-365.
- Hon C-Y, Peters CE, Jardine KJ, Arrandale VH. 2016. Historical occupational isocyanate exposure levels in two Canadian provinces. *J Occup Environ Hyg*: 00-00.
- INSEE. 2008. French classification of activities - NAF Rev. 2, 2008 (second edition, 2015).
- INRS. MetroPol.
- Kauffer E, Vincent R. 2007. Occupational Exposure to Mineral Fibres: Analysis of Results Stored on Colchic Database. *Ann Occup Hyg* 51: 131-142.

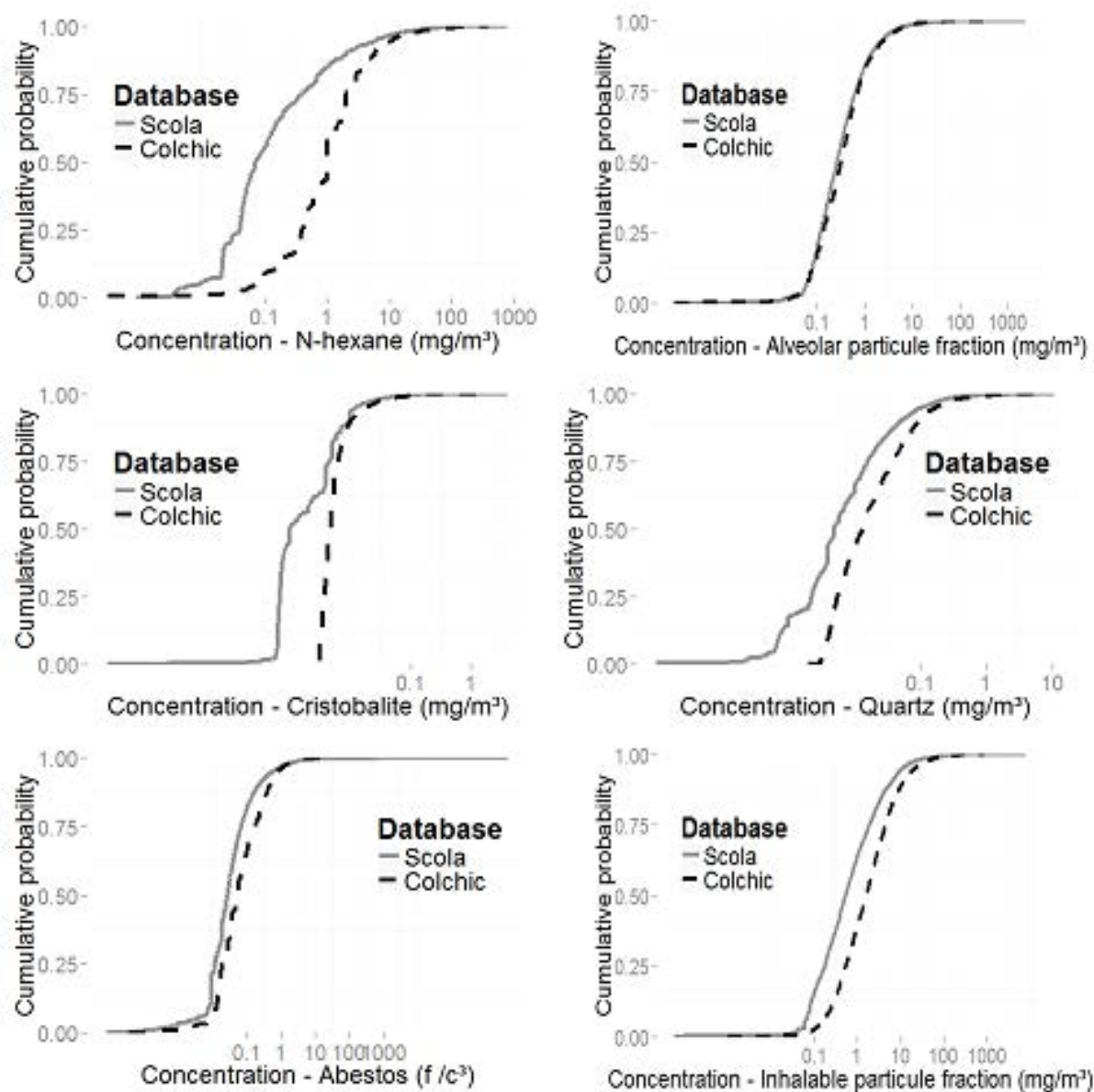
- Koh D-H, Bhatti P, Coble JB, Stewart PA, Lu W, Shu X-O, Ji B-T, Xue S, Locke SJ, Portengen L, Yang G, Chow W-H, Gao Y-T, Rothman N, Vermeulen R, Friesen MC. 2014. Calibrating a population-based job-exposure matrix using inspection measurements to estimate historical occupational exposure to lead for a population-based cohort in Shanghai, China. *J Expos Sci Environ Epidemiol* 24: 9-16.
- KOLSTAD HA, SØNDERSKOV J, BURSTYN I. 2005. Company-Level, Semi-Quantitative Assessment of Occupational Styrene Exposure when Individual Data are not Available. *Ann Occup Hyg* 49: 155-165.
- Koppisch D, Schinkel J, Gabriel S, Fransman W, Tielemans E. 2012. Use of the MEGA Exposure Database for the Validation of the Stoffenmanager Model. *Ann Occup Hyg* 56: 426-439.
- Krishnamoorthy K, Mallick A, Mathew T. 2009. Model-Based Imputation Approach for Data Analysis in the Presence of Non-detects. *Ann Occup Hyg* 53: 249-263.
- Lavoué J, BÉGIN D, BEAUDRY C, GÉRIN M. 2007. Monte Carlo Simulation to Reconstruct Formaldehyde Exposure Levels from Summary Parameters Reported in the Literature. *Ann Occup Hyg* 51: 161-172.
- Lavoue J, Droz PO. 2009. Multimodel Inference and Multimodel Averaging in Empirical Modeling of Occupational Exposure Levels. *Ann Occup Hyg* 53: 173-180.
- Lavoue J, Friesen MC, Burstyn I. 2012. Workplace Measurements by the US Occupational Safety and Health Administration since 1979: Descriptive Analysis and Potential Uses for Exposure Assessment. *Ann Occup Hyg*.
- Lavoué J, Gérin M, Vincent R. 2011. Comparison of Formaldehyde Exposure Levels in Two Multi-Industry Occupational Exposure Databanks Using Multimodel Inference. *J Occup Environ Hyg* 8: 38-48.
- Lavoué J, Vincent R, Gérin M. 2006. Statistical Modelling of Formaldehyde Occupational Exposure Levels in French Industries, 1986–2003. *Ann Occup Hyg* 50: 305-321.
- Le ministre du travail drs, de la famille, de la solidarité et de la ville. 2009. Décret n° 2009-1570 du 15 décembre 2009 relatif au contrôle du risque chimique sur les lieux de travail. *Journal officiel de la république française*. p 21758.

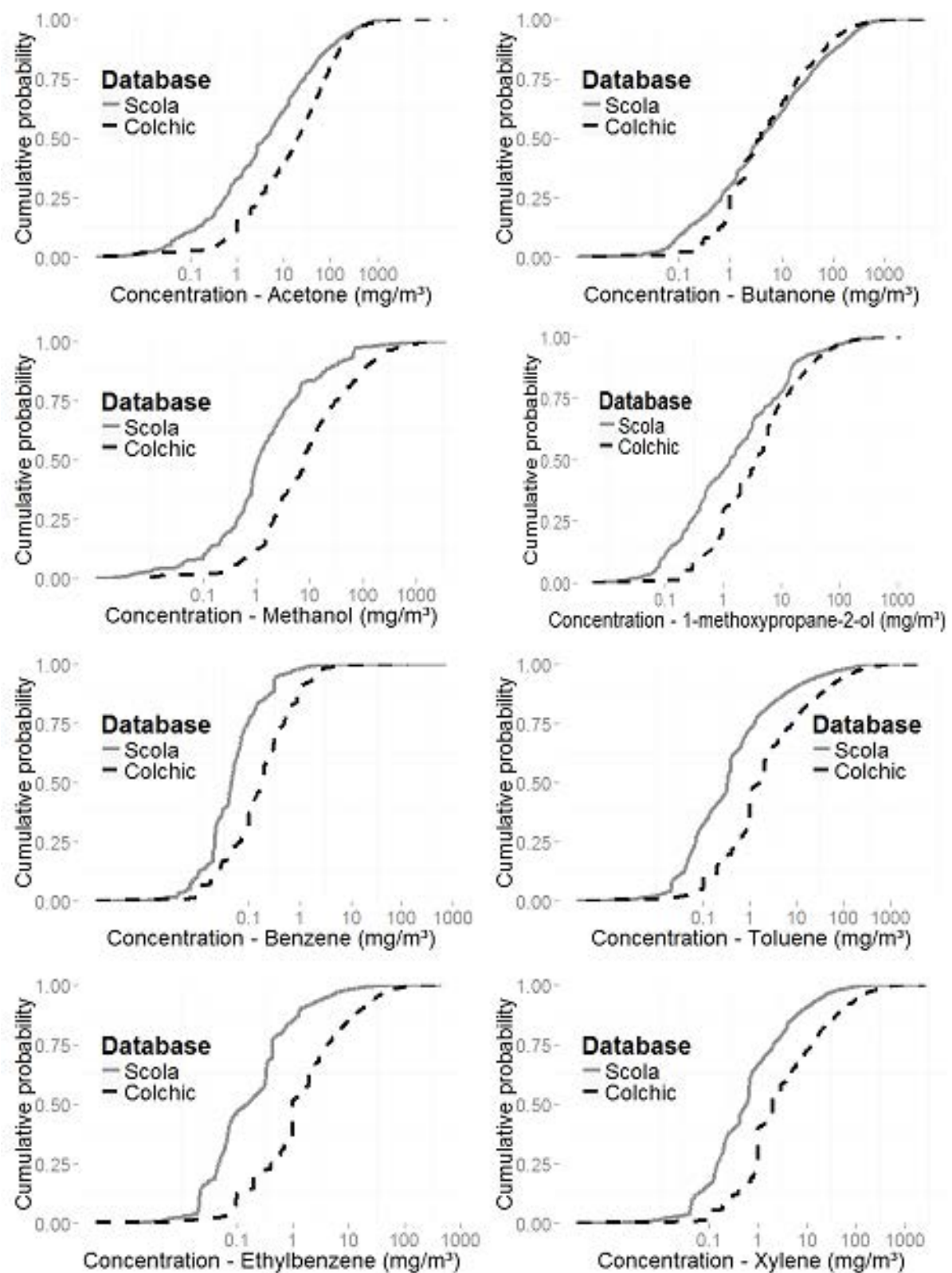
- Linch K, Miller W, Althouse R. 1998. Surveillance of Respirable Crystalline Silica Dust Using OSHA Compliance Data (1979-1995). *Am J Ind Med* 34: 547-558.
- Mater G, Paris C, Lavoué J. 2016. Descriptive analysis and comparison of two French occupational exposure databases: COLCHIC and SCOLA. *Am J Ind Med* 59: 1-13.
- Melville R, Lippmann M. 2001. Influence of Data Elements in OSHA Air Sampling Database on Occupational Exposure Levels. *Appl Occup Environ Hyg* 16: 884-899.
- Ogden TL. 2010. Handling results below the level of detection. *Ann Occup Hyg* 54: 255-256.
- Okun A, Cooper G, Bailer AJ, Bena J, Stayner L. 2004. Trends in occupational lead exposure since the 1978 OSHA lead standard. *Am J Ind Med* 45: 558-572.
- Olsson AC, Gustavsson P, Kromhout H, Peters S, Vermeulen R, Bröske I, Pesch B, Siemiatycki J, Pintos J, Brüning T, Cassidy A, Wichmann H-E, Consonni D, Landi MT, Caporaso N, Plato N, Merletti F, Mirabelli D, Richiardi L, Jöckel K-H, Ahrens W, Pohlabeln H, Lissowska J, Szeszenia-Dabrowska N, Zaridze D, Stücker I, Benhamou S, Bencko V, Foretova L, Janout V, Rudnai P, Fabianova E, Dumitru RS, Gross IM, Kendzia B, Forastiere F, Bueno-de-Mesquita B, Brennan P, Boffetta P, Straif K. 2011. Exposure to Diesel Motor Exhaust and Lung Cancer Risk in a Pooled Analysis from Case-Control Studies in Europe and Canada. *Am J Resp Crit Care Med* 183: 941-948.
- Pesch B, Kendzia B, Hauptmann K, Van Gelder R, Stamm R, Hahn J-U, Zschiesche W, Behrens T, Weiss T, Siemiatycki J, Lavoué J, Jöckel K-H, Brüning T. 2015. Airborne exposure to inhalable hexavalent chromium in welders and other occupations: Estimates from the German MEGA database. *Int J Hyg Env Health* 218: 500-506.
- Peters S, Vermeulen R, Portengen L, Olsson A, Kendzia B, Vincent R, Savary B, Lavoué J, Cavallo D, Cattaneo A, Mirabelli D, Plato N, Fevotte J, Pesch B, Brüning T, Straif K, Kromhout H. 2011. Modelling of occupational respirable crystalline silica exposure for quantitative exposure assessment in community-based case-control studies. *J Environ Monit* 13: 3262-3268.
- Pôle Emploi. 1999. Répertoire Opérationnel des Métiers et des Emplois.

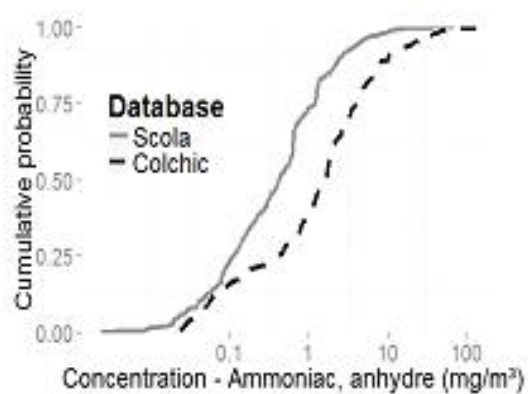
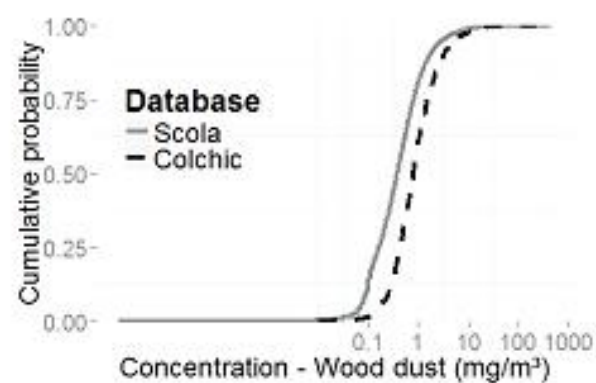
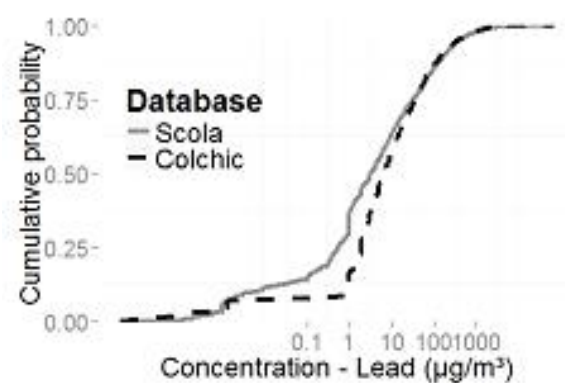
- Raaschou-Nielsen O, Hansen J, Thomsen BL, Johansen I, Lipworth L, McLaughlin JK, Olsen JH. 2002. Exposure of Danish Workers to Trichloroethylene, 1947-1989. *Appl Occup Environ Hyg* 17: 693-703.
- Rajan B, Alesbury R, Carton B, Gérin M, Litske H, Marquart H, Olsen E, Scheffers T, Stamm R, Woldbaek T. 1997. European Proposal for Core Information for the Storage and Exchange of Workplace Exposure Measurements on Chemical Agents. *Appl Occup Environ Hyg* 12: 31-39.
- Sarazin P, Burstyn I, Kincl L, Lavoué J. 2016. Trends in OSHA Compliance Monitoring Data 1979–2011: Statistical Modeling of Ancillary Information across 77 Chemicals. *Ann Occup Hyg*: 1-21.
- Sauvé J, Beaudry C, Bégin D, Dion C, Gérin M, Lavoué J. 2012. Silica Exposure During Construction Activities: Statistical Modeling of Task-Based Measurements from the Literature. *Ann Occup Hyg* 57: 432-443.
- Scarselli A. 2011. The National Registry of Occupational Exposures to Carcinogens (SIREP): information system and results. *Giornale italiano di medicina del lavoro ed ergonomia* 33: 78-79.
- Scarselli A, Binazzi A, Di Marzio D, Marinaccio A, Iavicoli S. 2012. Hexavalent Chromium Compounds in the Workplace: Assessing the Extent and Magnitude of Occupational Exposure in Italy. *J Occup Environ Hyg* 9: 398-407.
- Scarselli A, Montaruli C, Marinaccio A. 2007. The Italian Information System on Occupational Exposure to Carcinogens (SIREP): Structure, Contents and Future Perspectives. *Ann Occup Hyg* 51: 471-478.
- Stamm R. 2001. MEGA-Database: One Million Data Since 1972. *Appl Occup Environ Hyg* 16: 159-163.
- Stewart P, Rice C. 1990. A source of Exposure Data for Occupational epidemiology Studies. *Appl Occup Environ Hyg* 5: 359-363.
- Symanski E, Kupper LL, Hertz-Picciotto I, Rappaport SM. 1998. Comprehensive evaluation of long-term trends in occupational exposure: Part 2. Predictive models for declining exposures. *Occup Environ Med* 55: 310-316.

- Symanski E, Kupper LL, Rappaport SM. 1998. Comprehensive evaluation of long-term trends in occupational exposure: Part 1. Description of the database. *Occup Environ Med* 55: 300-309.
- Symanski E, Sällsten G, Chan W, Barregård L. 2001. Heterogeneity in sources of exposure variability among groups of workers exposed to inorganic mercury. *Ann Occup Hyg* 45: 677-687.
- Tobin J. 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica* 26: 24-36.
- Valiante DJ, Richards TB, Kinsley KB. 1992. Silicosis surveillance in New Jersey: Targeting workplaces using occupational disease and exposure surveillance data. *Am J Ind Med* 21: 517-526.
- van Houwelingen HC, Arends LR, Stijnen T. 2002. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 21: 589-624.
- Viechtbauer W. 2010. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 36: 1-48.
- Vincent R, Jeandel B. 2001. COLCHIC - Occupational Exposure to Chemical Agents Database: Current Content and Development Perspectives. *Appl Occup Environ Hyg* 16: 115-121.
- Vincent R, Jeandel B. 2002. Occupational exposure to lead. Information issued from the Colchic database. *Hygiène et sécurité du travail* 187: 63-72.
- Vinzents P, Carton B, Fjeldstad P. 1994. Extraction of Core Information and Possibilities for Comparison Between European Databases for Occupational Air Pollution Measurements. *Appl Occup Environ Hyg* 10(4): 351-354.

ANNEXE IV.1: MEASUREMENT DISTRIBUTION CURVES BY DATABASE FOR THE PERIOD 2007 – 2012 (COURBES DE DISTRIBUTION CUMULEE DES EXPOSITIONS PAR AGENT CHIMIQUE POUR LA PERIODE 2007 A 2012 POUR COLCHIC ET SCOLA)







ANNEXE IV.2 : COMPARISON OF EXPOSURE LEVELS IN CLASSES OF ACTIVITY COMMON TO COLCHIC AND SCOLA OVER THE PERIOD 2007 TO 2012 RESTRICTED TO THE SAMPLE DURATION BETWEEN 240 AND 480 MINUTES (COMPARAISON DES NIVEAUX D'EXPOSITION PAR CLASSE D'ACTIVITE COMMUNE A COLCHIC ET SCOLA SUR LA PÉRIODE 2007 À 2012 ET RESTREINTE AUX DURÉES DE PRÉLÈVEMENT COMPRISES ENTRE 240 ET 480 MINUTES)

Industrial sector

Chemical agent	N ¹	% included SCOLA ²	% included COLCHIC ³	Exceedance			Geometric Mean			Non-detects		
				Bias ⁴	Median SCOLA ⁵	Median COLCHIC ⁶	Bias GM ⁷	Median GM SCOLA ⁸	Median GM COLCHIC ⁹	Bias ND ¹⁰	Median ND SCOLA ¹¹	Median ND COLCHIC ¹²
Lead	10	72	80	2.4	19	38	-151	0.46	0.82	18.9	9.2	0
Wood (dust)	15	87	96	21.3	8.7	12	48.3	1.8	5.5	-11.8	22	31

Occupation

Chemical agent	N ¹	% included SCOLA ²	% included COLCHIC ³	Exceedance			Geometric Mean			Non-detects		
				Bias ⁴	Median SCOLA ⁵	Median COLCHIC ⁶	Bias GM ⁷	Median GM SCOLA ⁸	Median GM COLCHIC ⁹	Bias ND ¹⁰	Median ND SCOLA ¹¹	Median ND COLCHIC ¹²
Cristobalite	16	50	72	0.38	1.2	0.23	-2,34	0.0012	0.0027	-0.18	94	94
Wood (dust)	11	71	94	7.22	24	33	37	0.47	0.79	-8.03	7.5	0.67
Quartz	16	49	71	5.22	4.5	9.2	33	0.006	0.013	-15.2	48	25

Task

Chemical agent	N ¹	% included SCOLA ²	% included COLCHIC ³	Exceedance			Geometric Mean			Non-detects		
				Bias ⁴	Median SCOLA ⁵	Median COLCHIC ⁶	Bias GM ⁷	Median GM SCOLA ⁸	Median GM COLCHIC ⁹	Bias ND ¹⁰	Median ND SCOLA ¹¹	Median ND COLCHIC ¹²
Cristobalite	16	32	44	0.39	0	0	-646	0.0008	0.0026	-4.92	95	94
Quartz	16	31	44	5.15	5.2	6.1	43.6	0.004	0.016	-19.2	48	31

1: Number of classes categories in common with more than 20 measures in the two data

2: Percentage of SCOLA records included in the comparison (%)

3: Percentage of COLCHIC records included in the comparison (%)

4: Median absolute difference in exceedance fraction of OELV (%)

5: Median of exceedance fraction of OELV across activity categories in SCOLA (%)

6: Median of exceedance fraction of OELV across activity categories in COLCHIC (%)

7: Median relative difference in geometric mean (%)

8: Median of geometric mean values across activity categories in SCOLA (mg/m³ except lead in µg/m³)

9: Median of geometric mean values across activity categories in COLCHIC (mg/m³ except lead in µg/m³)

10: Median absolute difference in proportion of records reported as non detected (%)

11: Median of proportion of records reported as non detected across activity categories in SCOLA (%)

12: Median of proportion of records reported as non detected across activity categories in COLCHIC (%)

ANNEXE V.1 POURCENTAGE DE DONNÉES INFÉRIEURES À LA LIMITE DE QUANTIFICATION, RÉPARTITION DE LA PROPORTION DE MESURES COURTE DURÉE (CD) ET LONGUE DURÉE (LD) PAR AGENT CHIMIQUE COMPTANT PLUS DE 1 000 RÉSULTATS SUR LA PÉRIODE 1987-2015 POUR LES MESURES INDIVIDUELLES ET LES MESURES D'AMBIANCE DE TRAVAIL

Agent chimique 1987 2015	Individuel				Ambiant			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)
Poussières inhalables (A)	41686	4,3	7	93	25152	12	5	95
Acétone (S)	22914	9,5	30	70	9922	13	31	69
Toluène (S)	21509	20	36	64	11627	27	32	68
Fer (M)	16054	7	10	90	8147	13	8	92
Styrène (A)	15366	5,1	29	71	6149	11	29	71
Xylène (S)	15089	25	38	62	8361	24	32	68
Butanone (S)	14939	15	33	67	6805	21	29	71
Plomb (M)	14625	24	9	91	8985	35	8	92
Acétate d'éthyle (S)	13610	15	32	68	6409	16	30	70
Poussières alvéolaires (A)	12748	5,1	1	99	11956	10	1	99
Hydrocarbures aliphatiques et alicycliques (S)	11929	16	33	67	5945	16	25	75
Manganèse (M)	11327	14	10	90	5132	25	8	92
Chrome total (M)	10525	37	10	90	6676	51	9	91
Bois poussières de (A)	10148	2,7	2	98	4899	10	2	98
Ethylbenzène (S)	9781	26	40	60	5496	28	30	70
Acétate de n-butyle (S)	9152	17	36	64	4697	19	32	68
Nickel (M)	8987	35	7	93	5123	48	6	94
Formaldéhyde (A)	8593	4	37	63	10094	5	25	75
Ethanol (S)	8556	20	31	69	5456	16	29	71
Cuivre (M)	7450	23	6	94	3943	35	5	95
Quartz (A)	7351	38	1	99	5444	54	1	99
Alcool isopropylique (S)	7304	15	35	65	4330	15	32	68
Zinc (M)	7168	13	9	91	4189	24	6	94
Fibres Lsup5µm Dinf3µm (A)	6706	17	37	63	4845	34	25	75
Cristobalite (A)	6544	92	1	99	5340	94	1	99
Cobalt (M)	6001	42	4	96	3863	54	6	94
4-Méthylpentane-2-one (S)	5598	23	35	65	2951	26	30	70
Fibres Lsup5µm Dsup3µm (A)	5302	56	32	68	3552	80	20	80
Dichlorométhane (S)	5106	18	34	66	2705	22	34	66
n-Hexane (S)	4915	33	40	60	2400	39	32	68
Hydrocarbures aromatiques monocycliques (S)	4726	19	30	70	3098	19	20	80
1-Méthoxypropane-2-ol (S)	4527	23	29	71	2506	23	25	75
Aluminium (M)	4377	31	6	94	2391	47	8	92
Acétate de 2-méthoxy-1-méthyléthyle (S)	4161	21	33	67	2143	21	27	73
Benzène (S)	3742	62	39	61	2499	68	29	71

Agent chimique 1987 2015	Individuel				Ambiant			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)
Méthacrylate de méthyle (S)	3575	43	27	73	1459	38	21	79
Cadmium (M)	3343	61	9	91	2153	67	7	93
Trichloréthylène (S)	3033	15	40	60	2427	15	40	60
Benzo[a]pyrène (A)	2866	53	6	94	3119	52	8	92
Alcool butylique (S)	2798	25	32	68	1643	25	27	73
Acétate d'isopropyle (S)	2795	10	33	67	1214	12	34	66
Cyclohexanone (S)	2793	29	36	64	1610	33	30	70
Tétrachloroéthylène (S)	2715	15	35	65	2491	10	36	64
Poussières totales (A)	2702	3,9	7	93	5988	6	15	85
2-Méthylpropane-1-ol (S)	2571	21	34	66	1515	24	33	67
Triméthylbenzène (S)	2326	15	44	56	1003	17	47	53
Chrome VI soluble (M)	2208	49	10	90	1455	57	10	90
Acétate d'isobutyle (S)	2132	18	39	61	-	-	-	-
Titane (M)	1932	20	8	92	-	-	-	-
Etain (M)	1894	56	9	91	1154	70	5	95
Cyclohexane (S)	1878	20	32	68	1132	34	25	75
Acétate de 2-éthoxyéthyle (S)	1754	25	37	63	1006	21	35	65
Hydrocarbures benzéniques C9 à C12 (S)	1743	19	34	66	-	-	-	-
1,1,1-Trichloroéthane (S)	1739	11	51	49	1125	14	44	56
Hydrocarbures C6 a C12 (S)	1589	13	33	67	-	-	-	-
Béryllium (M)	1568	72	4	96	1099	71	4	97
Chrome VI insoluble (M)	1505	31	4	96	-	-	-	-
4-Hydroxy, 4-méthylpentan-2-one (S)	1492	34	33	67	-	-	-	-
Méthanol (S)	1387	37	35	65	-	-	-	-
Benzo[k]fluoranthène (A)	1380	43	7	93	1225	43	3	97
2-Butoxyéthanol (S)	1363	46	27	73	-	-	-	-
Benzoeacephenanthrylène (A)	1333	43	7	93	1171	43	3	97
Aérosol inhalable de fluides d'usinage (A)	1214	6,3	0	100	-	-	-	-
Acétate de propyle (S)	1187	22	31	69	-	-	-	-
Molybdène (M)	1170	58	5	95	-	-	-	-
Benzo[ghi]pérylène (A)	1166	52	6	94	1013	55	3	97
Tétrahydrofuranne (S)	1161	28	29	71	-	-	-	-
2-Phénylpropène (A)	1154	21	27	73	-	-	-	-
Benzo[a]anthracène (A)	1121	41	6	94	-	-	-	-
n-Heptane (S)	1026	26	42	58	-	-	-	-
Equivalent NCl3 (A)	-	-	-	-	3479	5	2	98
Fraction soluble CH2 Cl2 (A)	-	-	-	-	2027	17	8	92
Mercure (M)	-	-	-	-	1652	26	17	83
Particules diesel carbone élémentaire (A)	-	-	-	-	1609	22	1	99
Particules diesel carbone organique (A)	-	-	-	-	1609	21	1	99

Agent chimique 1987 2015	Individuel				Ambiant			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)
Poussières thoraciques (A)	-	-	-	-	1554	5	0	100
Diméthylnitrosoamine (A)	-	-	-	-	1476	17	0	100
Nitrosamines totales (A)	-	-	-	-	1314	14	0	100
Acétaldéhyde (A)	-	-	-	-	1293	14	24	76
Eau oxygénée (A)	-	-	-	-	1246	20	29	71
Oxyde de diazote (A)	-	-	-	-	1173	8	97	3
4,4-Diisocyanate de diphenylméthane (A)	-	-	-	-	1160	78	31	69
Ammoniac anhydre (A)	-	-	-	-	1097	5	17	83
Chlorures gazeux (A)	-	-	-	-	1069	33	11	89
Chlorures particuliers (A)	-	-	-	-	1048	29	11	89
Glutaral (A)	-	-	-	-	1035	31	36	64

Famille des agents chimiques : (A) Autre famille, (M) Métal et (S) Solvant

ANNEXE V.2 : POURCENTAGE DE DONNÉES INFÉRIEURES À LA LIMITE DE QUANTIFICATION, RÉPARTITION DE LA PROPORTION DE MESURES COURTE DURÉE (CD) ET LONGUE DURÉE (LD) PAR AGENT CHIMIQUE COMPTANT PLUS DE 1 000 RÉSULTATS SUR LA PÉRIODE 2002-2015 POUR LES MESURES INDIVIDUELLES ET LES MESURES D'AMBIANCE DE TRAVAIL

Agent chimique 2002 2015	Individuel				Ambiant			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)
Poussières inhalables (A)	18474	3,6	4	96	11007	11	3	97
Acétone (S)	10033	13	28	72	4288	16	18	82
Fer (M)	8331	9,6	5	95	3486	17	3	97
Toluène (S)	8148	28	29	71	4586	33	19	81
Bois poussières de (A)	7229	1,8	2	98	3350	9	3	97
Styrène (A)	6950	7,5	28	72	2900	13	16	84
Xylène (S)	6432	31	31	69	3610	26	19	81
Manganèse	6415	14	5	95	2485	26	3	97
Butanone (S)	6254	21	26	74	2775	25	16	84
Hydrocarbures aliphatiques et alicycliques (S)	6165	21	28	72	3155	19	18	82
Acétate d'éthyle (S)	5416	20	26	74	2371	20	16	84
Formaldéhyde (A)	5340	3,7	30	70	5989	5	18	82
Chrome total (M)	4906	36	5	95	2530	53	4	96
Ethylbenzène (S)	4867	34	33	67	2931	35	19	81
Ethanol (S)	4540	21	26	74	2717	17	19	81
Nickel (M)	4486	34	5	95	2107	50	2	98
Plomb (M)	4419	44	6	94	2634	55	4	96
Cuivre (M)	4330	24	5	95	1779	39	3	97
Poussières alvéolaires (A)	4159	8,7	1	99	3070	15	1	99
Zinc (M)	3935	13	6	94	1677	26	3	97
Quartz (A)	3855	35	1	99	2701	50	1	99
Cristobalite (A)	3843	93	1	99	2705	93	1	99
Acétate de n-butyle (S)	3653	24	31	69	1942	27	21	79
Alcool isopropylique (S)	3516	20	31	69	1769	20	23	77
Aluminium (M)	2912	30	4	96	1252	43	4	96
Cobalt (M)	2678	51	3	97	1555	65	3	97
Méthacrylate de méthyle (S)	2468	49	27	73	-	-	-	-
n-Hexane (S)	2301	37	29	71	1319	49	20	80
4-Méthylpentane-2-one (S)	2265	33	32	68	1131	38	20	80
1-Méthoxypropane-2-ol (S)	2235	29	23	77	1136	31	16	84
Acétate de 2-méthoxy-1-méthyléthyle (S)	2204	26	27	73	1091	27	20	80
Hydrocarbures aromatiques monocycliques (S)	1936	27	26	74	1209	19	8	92
Dichlorométhane (S)	1656	31	23	77	-	-	-	-
Cadmium (M)	1632	73	4	96	-	-	-	-
Benzo[a]pyrène (A)	1453	59	7	93	1207	65	4	96

Agent chimique 2002 2015	Individuel				Ambiant			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)
Fibres Lsup5µm Dinf3µm (A)	1434	30	27	73	1044	54	12	88
Tétrachloroéthylène (S)	1410	18	28	72	1320	12	21	79
Benzène (S)	1405	69	28	72	1096	74	12	88
Fibres Lsup5µm Dsup3µm (A)	1393	74	27	73	-	-	-	-
Titane (M)	1362	17	7	93	-	-	-	-
Chrome VI soluble (M)	1306	35	4	96	-	-	-	-
Chrome VI insoluble (M)	1200	28	4	96	-	-	-	-
Aérosol inhalable de fluides d'usinage (A)	1200	6,4	0	100	-	-	-	-
Hydrocarbures benzéniques C9 à C12 (S)	1153	22	36	64	-	-	-	-
Hydrocarbures C6 a C12 (S)	1103	16	30	70	-	-	-	-
Alcool butylique (S)	1098	30	27	73	-	-	-	-
Cyclohexanone (S)	1030	34	25	75	-	-	-	-
Equivalent NCl3 (A)	-	-	-	-	1307	9	3	97
Particules diesel carbone élémentaire (A)	-	-	-	-	1221	27	1	99
Particules diesel carbone organique (A)	-	-	-	-	1221	26	1	99

Famille des agents chimiques : (A) Autre famille, (M) Métal et (S) Solvant

**ANNEXE V.3 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE)
ET DE LEUR INTERVALLE DE CONFIANCE À 95 % (IC 95 %) POUR LES
PRÉLÈVEMENTS INDIVIDUELS PAR MÉTA ANALYSE STRATIFIÉE PAR
PÉRIODE 1987-2015 ET 2002-2015 ET PAR FAMILLE D'AGENTS CHIMIQUES**

Variable	Période 1987-2015		Période 2002-2015	
Catégorie	Solvant N=35 RIE (IC 95 %)	Métaux N=13 RIE (IC 95 %)	Solvant N=20 RIE (IC 95 %)	Métaux N=12 RIE (IC 95 %)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	34 % [29-39]	37 % [29-44]	37 % [30-42]	_(2)
Longue durée (≥ 60 minutes)	3 % [2-4]	5 % [3-7]	2 % [1-4]	_(2)
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,57 (0,54;0,60)	0,51 (0,44;0,61)	0,55 (0,51;0,59)	0,56 (0,50;0,62)
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	5 % [4-6]	7 % [3-10]	10 % [8-12]	_(2)
Longue durée (≥ 60 minutes)	5 % [4-6]	4 % [2-6]	10 % [9-12]	7 % [3-10]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,07 (0,91;1,26)	1,17 (0,96;1,44)	1,15 (0,96;1,37)	1,38 (1,10;1,72)
Bourgo_comte	1,01 (0,87;1,18)	1,07 (0,85;1,33)	1,12 (0,91;1,36)	1,17 (0,95;1,43)
Bretagne	0,84 (0,67;1,06)	1,12 (0,84;1,50)	0,71 (0,46;1,11)	1,10 (0,84;1,44)
CALA	1,03 (0,90;1,18)	1,25 (0,97;1,62)	1,28 (1,04;1,58)	1,43 (1,14;1,79)
Centre	1,04 (0,88;1,22)	1,15 (0,81;1,63)	1,17 (0,89;1,53)	0,97 (0,73;1,29)
DOM Corse	0,87 (0,71;1,06)	0,94 (0,70;1,27)	1,04 (0,79;1,38)	0,94 (0,71;1,24)
IDF	0,99 (0,86;1,14)	1,01 (0,78;1,32)	0,89 (0,71;1,12)	1,17 (0,97;1,41)
MPLR	1,00 (0,88;1,15)	1,28 (1,05;1,57)	1,09 (0,90;1,32)	1,19 (0,99;1,43)
Nord	1,12 (0,92;1,36)	1,23 (0,93;1,63)	1,45 (1,06;1,97)	1,43 (0,98;2,10)
Normandie	1,29 (1,13;1,48)	1,25 (0,96;1,63)	1,25 (0,95;1,65)	1,04 (0,79;1,36)
P_Loire	0,94 (0,79;1,13)	0,80 (0,55;1,18)	1,03 (0,76;1,39)	0,71 (0,55;0,91)
PACA	1,23 (1,01;1,50)	1,55 (1,06;2,27)	1,35 (1,02;1,79)	1,08 (0,74;1,58)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Autre raison	0,95 (0,88;1,02)	0,98 (0,85;1,14)	0,84 (0,73;0,96)	0,73 (0,60;0,89)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Entreprise	0,89 (0,83;0,95)	1,11 (0,89;1,39)	0,86 (0,77;0,96)	0,91 (0,70;1,19)
Corps médical	1,11 (0,99;1,23)	0,99 (0,74;1,32)	1,31 (1,14;1,49)	0,66 (0,48;0,91)
Autre origine	0,63 (0,40;1,00)	1,12 (0,66;1,92)	0,88 (0,43;1,81)	0,61 (0,46;0,83)

Variable	Période 1987–2015		Période 2002-2015	
	Solvant N=35 RIE (IC 95 %)	Métaux N=13 RIE (IC 95 %)	Solvant N=20 RIE (IC 95 %)	Métaux N=12 RIE (IC 95 %)
Ventilation générale				
Non mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,09 (1,01;1,18)	1,02 (0,88;1,18)	1,02 (0,91;1,13)	0,85 (0,69;1,04)
Oui	1,20 (1,11;1,30)	0,97 (0,85;1,12)	1,07 (0,93;1,23)	1,06 (0,90;1,24)
Non	1,07 (1,00;1,14)	1,08 (0,93;1,24)	0,90 (0,80;1,03)	1,00 (0,85;1,18)
Information non relevée	0,94 (0,82;1,06)	1,06 (0,87;1,30)	0,90 (0,77;1,05)	0,81 (0,70;0,94)
Tranche d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,15 (1,03;1,28)	1,38 (1,14;1,68)	1,11 (0,96;1,28)	1,18 (0,90;1,56)
50-199 salariés	1,04 (0,93;1,17)	1,34 (1,12;1,61)	1,03 (0,88;1,21)	1,19 (0,90;1,58)
200 salariés ou plus	1,22 (1,10;1,35)	1,61 (1,27;2,05)	1,10 (0,93;1,31)	1,09 (0,69;1,72)
Information non relevée	1,11 (0,98;1,26)	1,34 (1,12;1,60)	1,08 (0,90;1,30)	1,08 (0,88;1,33)
Équipement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,54 (1,41;1,68)	1,97 (1,67;2,33)	1,72 (1,52;1,94)	2,25 (1,85;2,75)
Information non relevée	0,95 (0,84;1,08)	0,92 (0,74;1,14)	0,99 (0,78;1,27)	0,98 (0,82;1,17)
Type de procédé				
Ouvert	-	-	1,00 (Réf.)	1,00 (Réf.)
Semi-ouvert	-	-	0,73 (0,63;0,84)	0,52 (0,40;0,68)
Fermé	-	-	0,75 (0,64;0,87)	0,55 (0,43;0,70)
Autre type de procédé	-	-	0,87 (0,69;1,09)	0,50 (0,25;1,01)
Fréquence d'exposition				
Permanent	-	-	1,00 (Réf.)	1,00 (Réf.)
Fréquent	-	-	1,07 (0,98;1,18)	1,05 (0,97;1,14)
Occasionnel	-	-	0,96 (0,87;1,07)	0,70 (0,60;0,83)
Autre fréquence d'exposition	-	-	0,91 (0,73;1,13)	0,84 (0,43;1,63)

⁽¹⁾Données calculées pour les solvants et pour les métaux respectivement avec 30 et 6 agents chimiques sur 1987 – 2015 et avec 20 et 1 agents chimiques sur 2002 – 2015.

⁽²⁾Pour les métaux sur 2002-2015, un seul agent chimique appartient à la famille. L'analyse porte donc uniquement sur les mesures de « longue durée ».

**ANNEXE V.4 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE)
ET DE LEUR INTERVALLE DE CONFIANCE À 95 % (IC 95 %) POUR LES
PRÉLÈVEMENTS D'AMBIANCE PAR MÉTA ANALYSE STRATIFIÉE PAR
PÉRIODE 1987 - 2015 ET 2002 – 2015 ET PAR FAMILLE D'AGENTS CHIMIQUES**

Variable	Période 1987 – 2015		Période 2002-2015	
Catégorie	Solvant N=25 RIE (IC 95 %)	Métaux N=10 RIE (IC 95 %)	Solvant N=15 RIE (IC 95 %)	Métaux N=7 RIE (IC 95 %)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	30 % [14-43]	-	40 % [-72-79]	-
Longue durée (≥ 60 minutes)	4 % [3-5]	5 % [3-7]	4 % [3-6]	7 % [6-9]
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,56 (0,47;0,67)	-	0,49 (0,26;0,95)	-
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	5 % [2-7]	-	10 % [-3-21]	-
Longue durée (≥ 60 minutes)	4 % [2-7]	5 % [2-8]	10 % [7-13]	8 % [4-11]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,06 (0,84;1,33)	1,03 (0,76;1,40)	0,95 (0,73;1,22)	1,19 (0,80;1,78)
Bourgo_comte	0,97 (0,80;1,17)	1,16 (0,91;1,49)	1,02 (0,77;1,35)	1,22 (0,86;1,74)
Bretagne	0,77 (0,55;1,07)	0,76 (0,45;1,26)	0,70 (0,49;1,02)	0,97 (0,64;1,47)
CALA	1,03 (0,85;1,26)	1,24 (0,95;1,62)	1,27 (0,95;1,70)	1,64 (1,11;2,42)
Centre	1,11 (0,89;1,38)	1,08 (0,72;1,63)	1,00 (0,70;1,43)	1,29 (0,79;2,12)
DOM Corse	0,89 (0,68;1,17)	1,05 (0,60;1,82)	0,84 (0,62;1,14)	0,98 (0,70;1,38)
IDF	0,94 (0,75;1,17)	0,91 (0,72;1,15)	0,85 (0,67;1,10)	1,10 (0,80;1,50)
MPLR	1,08 (0,92;1,28)	1,10 (0,82;1,48)	0,99 (0,77;1,26)	1,11 (0,80;1,55)
Nord	1,34 (0,96;1,87)	1,07 (0,64;1,79)	1,34 (0,96;1,86)	0,51 (0,30;0,88)
Normandie	1,31 (1,03;1,67)	1,17 (0,88;1,55)	1,02 (0,75;1,40)	0,82 (0,54;1,25)
P_Loire	0,88 (0,71;1,10)	0,81 (0,48;1,37)	0,90 (0,66;1,23)	0,69 (0,46;1,03)
PACA	0,97 (0,76;1,24)	1,53 (1,11;2,10)	1,11 (0,67;1,84)	0,90 (0,57;1,40)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Autre raison	1,02 (0,91;1,15)	1,19 (1,05;1,35)	0,86 (0,65;1,14)	0,93 (0,77;1,11)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Entreprise	0,85 (0,76;0,96)	1,10 (0,91;1,33)	0,91 (0,77;1,07)	0,81 (0,65;1,01)
Corps médical	1,00 (0,89;1,13)	0,83 (0,67;1,03)	1,15 (0,94;1,41)	0,61 (0,47;0,79)
Autre origine	0,93 (0,47;1,85)	0,98 (0,66;1,47)	3,08 (0,50;18,81)	0,45 (0,25;0,80)

Variable	Période 1987 – 2015		Période 2002-2015	
	Solvant N=25 RIE (IC 95 %)	Métaux N=10 RIE (IC 95 %)	Solvant N=15 RIE (IC 95 %)	Métaux N=7 RIE (IC 95 %)
Ventilation générale				
Non mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,21 (1,02;1,44)	0,93 (0,62;1,40)	1,21 (1,00;1,47)	0,78 (0,62;0,98)
Oui	1,24 (1,01;1,51)	1,09 (0,77;1,55)	1,10 (0,87;1,40)	0,84 (0,66;1,06)
Non	0,97 (0,88;1,05)	0,89 (0,77;1,05)	0,66 (0,56;0,77)	1,02 (0,88;1,18)
Information non relevée	1,01 (0,85;1,20)	1,00 (0,85;1,18)	0,97 (0,78;1,19)	0,88 (0,63;1,23)
Tranche d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	0,88 (0,68;1,13)	1,24 (0,93;1,64)	1,07 (0,87;1,32)	1,19 (0,89;1,58)
50-199 salariés	0,93 (0,80;1,09)	1,05 (0,86;1,28)	0,75 (0,61;0,91)	1,00 (0,81;1,24)
200 salariés ou plus	0,98 (0,79;1,22)	1,01 (0,80;1,29)	0,71 (0,57;0,88)	1,04 (0,77;1,39)
Information non relevée	0,95 (0,78;1,16)	1,08 (0,84;1,40)	0,67 (0,52;0,85)	1,39 (1,02;1,91)
Équipement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,38 (1,18;1,61)	1,87 (1,48;2,38)	1,20 (1,00;1,45)	2,08 (1,65;2,63)
Information non relevée	1,00 (0,87;1,15)	0,94 (0,77;1,17)	0,79 (0,69;0,90)	0,73 (0,53;1,01)
Type de procédé				
Ouvert	-	-	1,00 (Réf.)	1,00 (Réf.)
Semi-ouvert	-	-	0,89 (0,75;1,05)	0,92 (0,73;1,16)
Fermé	-	-	1,03 (0,83;1,28)	0,72 (0,50;1,04)
Autre type de procédé	-	-	0,65 (0,42;1,01)	0,91 (0,46;1,81)
Fréquence d'exposition				
Permanent	-	-	1,00 (Réf.)	1,00 (Réf.)
Fréquent	-	-	1,10 (1,00;1,21)	0,99 (0,85;1,17)
Occasionnel	-	-	1,13 (0,95;1,34)	0,69 (0,56;0,85)
Autre fréquence d'exposition	-	-	0,82 (0,60;1,12)	0,55 (0,35;0,85)

⁽¹⁾Données calculées pour les solvants avec 14 agents chimiques sur 1987 – 2015 et 2 agents chimiques sur 2002 – 2015.
Pour les métaux, la proportion de prélèvements de « courte durée » est supérieure à 20 %. Les données ont été écartées. L'analyse porte donc uniquement sur les mesures de « longue durée ».

**ANNEXE V.5 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE)
ET DE LEUR INTERVALLE DE CONFIANCE À 95 % (IC 95 %) POUR LES
PRÉLÈVEMENTS INDIVIDUELS PAR MÉTA ANALYSE DES MODÈLES AVEC ET
SANS VARIABLE ALÉATOIRE STRATIFIÉE PAR PÉRIODE 1987-2015 ET 2002 –
2015**

Variable Catégorie	Période 1987–2015		Période 2002-2015	
	Aléatoire N=63 RIE (IC 95 %)	Sans aléatoire N=63 RIE (IC 95 %)	Aléatoire N=40 RIE (IC 95 %)	Sans aléatoire N=40 RIE (IC 95 %)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	35 % [30-39]	39 % [33-44]	37 % [31-42]	37 % [39-45]
Longue durée (≥ 60 minutes)	4 % [3-5]	4 % [3-5]	4 % [3-5]	4 % [3-5]
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,54 (0,49;0,59)	0,50 (0,45;0,55)	0,54 (0,50;0,58)	0,53 (0,49;0,58)
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	5 % [4-6]	6 % [5-7]	10 % [8-11]	10 % [8-12]
Longue durée (≥ 60 minutes)	5 % [4-6]	7 % [6-8]	8 % [6-10]	8 % [7-10]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,04 (0,95;1,13)	1,09 (0,99;1,20)	1,14 (1,01;1,29)	1,33 (1,18;1,49)
Bourgo_comte	0,97 (0,89;1,05)	0,94 (0,85;1,04)	1,00 (0,91;1,09)	1,03 (0,93;1,13)
Bretagne	0,94 (0,83;1,07)	0,92 (0,80;1,06)	0,86 (0,71;1,03)	0,70 (0,58;0,85)
CALA	1,00 (0,89;1,11)	0,88 (0,78;1,01)	1,14 (1,01;1,27)	1,13 (0,99;1,30)
Centre	1,02 (0,92;1,12)	1,07 (0,96;1,19)	1,08 (0,95;1,24)	1,16 (1,01;1,33)
DOM Corse	0,87 (0,78;0,98)	0,82 (0,71;0,94)	0,98 (0,87;1,11)	1,07 (0,92;1,24)
IDF	1,00 (0,92;1,08)	1,01 (0,92;1,11)	1,03 (0,93;1,14)	1,06 (0,96;1,17)
MPLR	0,99 (0,90;1,07)	0,95 (0,86;1,04)	1,04 (0,95;1,14)	1,13 (1,00;1,28)
Nord	1,06 (0,95;1,18)	1,03 (0,91;1,17)	1,29 (1,07;1,54)	1,33 (1,12;1,59)
Normandie	1,15 (1,04;1,27)	1,07 (0,94;1,22)	1,10 (0,95;1,26)	1,19 (1,01;1,41)
P_Loire	0,90 (0,80;1,00)	0,85 (0,76;0,96)	0,90 (0,78;1,05)	0,91 (0,77;1,06)
PACA	1,25 (1,08;1,44)	1,17 (1,02;1,34)	1,05 (0,87;1,26)	1,17 (0,98;1,39)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Autre raison	0,98 (0,93;1,03)	0,97 (0,91;1,03)	0,86 (0,78;0,95)	0,81 (0,73;0,89)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Entreprise	0,91 (0,87;0,97)	0,95 (0,90;1,01)	0,89 (0,82;0,98)	0,88 (0,81;0,95)
Corps médical	1,02 (0,94;1,11)	0,95 (0,88;1,02)	0,99 (0,87;1,12)	0,86 (0,77;0,97)
Autre origine	0,88 (0,71;1,11)	0,86 (0,69;1,07)	0,81 (0,63;1,04)	0,82 (0,59;1,13)

Variable	Période 1987–2015		Période 2002-2015	
	Aléatoire N=63 RIE (IC 95 %)	Sans aléatoire N=63 RIE (IC 95 %)	Aléatoire N=40 RIE (IC 95 %)	Sans aléatoire N=40 RIE (IC 95 %)
Ventilation générale				
Non mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,07 (1,00;1,13)	0,96 (0,90;1,02)	0,98 (0,92;1,05)	0,95 (0,88;1,04)
Oui	1,14 (1,08;1,21)	1,19 (1,11;1,28)	1,08 (1,00;1,17)	1,15 (1,04;1,27)
Non	1,06 (1,01;1,12)	1,11 (1,05;1,18)	0,97 (0,90;1,04)	0,98 (0,90;1,08)
Information non relevée	0,99 (0,91;1,07)	1,07 (0,98;1,18)	0,88 (0,80;0,98)	0,85 (0,75;0,96)
Tranche d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,16 (1,07;1,26)	1,11 (1,02;1,22)	1,10 (1,00;1,22)	1,02 (0,93;1,11)
50-199 salariés	1,07 (0,99;1,17)	1,06 (0,97;1,15)	1,04 (0,94;1,16)	1,01 (0,92;1,10)
200 salariés ou plus	1,14 (1,02;1,27)	1,06 (0,94;1,20)	1,05 (0,91;1,21)	1,02 (0,90;1,15)
Information non relevée	1,10 (1,02;1,20)	1,00 (0,92;1,07)	1,05 (0,95;1,17)	0,98 (0,89;1,08)
Équipement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,67 (1,57;1,77)	1,45 (1,32;1,59)	1,86 (1,71;2,01)	1,56 (1,40;1,74)
Information non relevée	0,99 (0,90;1,09)	0,87 (0,79;0,96)	0,96 (0,86;1,08)	0,92 (0,83;1,03)
Type de procédé				
Ouvert	-	-	1,00 (Réf.)	1,00 (Réf.)
Semi-ouvert	-	-	0,72 (0,66;0,79)	0,65 (0,58;0,71)
Fermé	-	-	0,73 (0,66;0,81)	0,65 (0,58;0,72)
Autre type de procédé	-	-	0,80 (0,65;0,99)	0,86 (0,66;1,11)
Fréquence d'exposition				
Permanent	-	-	1,00 (Réf.)	1,00 (Réf.)
Fréquent	-	-	1,06 (1,00;1,12)	0,94 (0,89;0,98)
Occasionnel	-	-	0,88 (0,82;0,95)	0,82 (0,75;0,90)
Autre fréquence d'exposition	-	-	0,89 (0,74;1,07)	1,10 (0,84;1,44)

**ANNEXE V.6 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE)
ET DE LEUR INTERVALLE DE CONFIANCE À 95 % (IC 95 %) POUR LES
PRÉLÈVEMENTS D'AMBIANCE PAR MÉTA ANALYSE DES MODÈLES AVEC ET
SANS VARIABLE ALÉATOIRE STRATIFIÉE PAR PÉRIODE 1987-2015 ET 2002-
2015**

Variable	Période 1987-2015		Période 2002-2015	
Catégorie	Aléatoire N=59 RIE (IC 95 %)	Sans aléatoire N=59 RIE (IC 95 %)	Aléatoire N=31 RIE (IC 95 %)	Sans aléatoire N=31 RIE (IC 95 %)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	30 % [17-41]	43 % [29-55]	40 % [-72-79]	45 % [-71-72]
Longue durée (≥ 60 minutes)	5 % [4-6]	4 % [4-5]	5 % [4-6]	4 % [3-5]
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,53 (0,46;0,63)	0,43 (0,35;0,52)	0,49 (0,26;0,95)	0,37 (0,20;0,70)
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	5 % [3-8]	3 % [0-7]	10 % [-3-21]	8 % [-3-17]
Longue durée (≥ 60 minutes)	4 % [3-5]	6 % [5-7]	8 % [6-9]	10 % [8-11]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,07 (0,97;1,19)	1,18 (1,05;1,34)	1,00 (0,90;1,12)	1,25 (1,06;1,47)
Bourgo_comte	1,00 (0,92;1,08)	1,06 (0,96;1,18)	0,98 (0,87;1,09)	1,15 (0,98;1,34)
Bretagne	0,92 (0,79;1,07)	0,83 (0,69;1,00)	0,99 (0,81;1,20)	0,95 (0,77;1,19)
CALA	1,08 (0,96;1,21)	1,10 (0,97;1,24)	1,11 (0,97;1,27)	1,28 (1,06;1,55)
Centre	1,06 (0,95;1,18)	1,07 (0,94;1,22)	1,11 (0,94;1,30)	1,14 (0,94;1,39)
DOM Corse	1,01 (0,87;1,16)	1,08 (0,92;1,27)	1,03 (0,86;1,24)	1,05 (0,87;1,28)
IDF	0,95 (0,86;1,06)	0,95 (0,85;1,07)	0,93 (0,84;1,03)	1,05 (0,91;1,21)
MPLR	1,03 (0,92;1,15)	0,95 (0,86;1,06)	1,09 (0,97;1,24)	1,14 (0,98;1,33)
Nord	1,10 (0,93;1,30)	1,18 (1,01;1,38)	1,09 (0,85;1,40)	1,11 (0,86;1,44)
Normandie	1,15 (1,04;1,27)	1,13 (0,99;1,29)	0,97 (0,85;1,11)	1,04 (0,90;1,22)
P_Loire	0,91 (0,79;1,04)	0,98 (0,86;1,12)	0,92 (0,78;1,08)	0,99 (0,83;1,19)
PACA	1,13 (1,02;1,25)	1,21 (1,04;1,40)	1,03 (0,85;1,25)	1,07 (0,87;1,32)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Autre raison	1,08 (1,01;1,17)	1,09 (1,01;1,17)	0,97 (0,87;1,09)	0,91 (0,81;1,03)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Entreprise	0,92 (0,86;0,98)	0,88 (0,83;0,94)	0,97 (0,89;1,05)	0,89 (0,82;0,97)
Corps médical	0,92 (0,85;1,00)	0,89 (0,82;0,96)	0,99 (0,88;1,12)	0,96 (0,85;1,09)
Autre origine	1,13 (0,92;1,37)	1,07 (0,86;1,32)	1,18 (0,70;2,02)	1,28 (0,73;2,23)

Variable	Période 1987-2015		Période 2002-2015	
	Aléatoire N=59 RIE (IC 95 %)	Sans aléatoire N=59 RIE (IC 95 %)	Aléatoire N=31 RIE (IC 95 %)	Sans aléatoire N=31 RIE (IC 95 %)
Ventilation générale				
Non mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,20 (1,08;1,33)	1,03 (0,94;1,12)	1,09 (0,98;1,22)	1,03 (0,94;1,13)
Oui	1,18 (1,06;1,32)	1,12 (1,02;1,23)	1,08 (0,94;1,24)	0,97 (0,85;1,10)
Non	0,98 (0,93;1,04)	1,12 (1,05;1,19)	0,84 (0,76;0,94)	0,88 (0,79;0,99)
Information non relevée	1,00 (0,91;1,10)	1,16 (1,04;1,29)	0,89 (0,78;1,02)	0,93 (0,80;1,09)
Tranche d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,00 (0,89;1,11)	0,90 (0,80;1,00)	1,11 (0,99;1,23)	1,11 (0,99;1,25)
50-199 salariés	0,97 (0,89;1,05)	0,89 (0,81;0,97)	0,96 (0,87;1,07)	0,97 (0,85;1,09)
200 salariés ou plus	0,98 (0,88;1,09)	0,94 (0,83;1,06)	0,93 (0,84;1,02)	0,85 (0,74;0,98)
Information non relevée	1,03 (0,94;1,13)	0,96 (0,88;1,05)	1,02 (0,87;1,18)	0,96 (0,83;1,11)
Équipement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,62 (1,46;1,81)	1,39 (1,25;1,55)	1,39 (1,22;1,59)	1,40 (1,21;1,61)
Information non relevée	1,04 (0,96;1,12)	0,91 (0,83;0,99)	0,89 (0,80;1,00)	0,76 (0,67;0,87)
Type de procédé				
Ouvert	-	-	1,00 (Réf.)	1,00 (Réf.)
Semi-ouvert	-	-	0,96 (0,88;1,06)	0,86 (0,78;0,95)
Fermé	-	-	0,92 (0,81;1,04)	0,80 (0,72;0,88)
Autre type de procédé	-	-	0,89 (0,73;1,09)	0,78 (0,60;1,02)
Fréquence d'exposition				
Permanent	-	-	1,00 (Réf.)	1,00 (Réf.)
Fréquent	-	-	1,06 (0,99;1,13)	1,03 (0,94;1,13)
Occasionnel	-	-	0,96 (0,86;1,08)	0,86 (0,78;0,95)
Autre fréquence d'exposition	-	-	0,78 (0,66;0,93)	0,92 (0,75;1,13)

**ANNEXE V.7 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE)
PAR MÉTA ANALYSE DES AGENTS MESURÉS EN INDIVIDUEL SUR UNE
LONGUE DURÉE (LD) ASSOCIÉE AUX PRÉLÈVEMENTS COURTE DURÉE SI
LEUR PROPORTION EST SUPÉRIEURE À 20 % (CD>20 %) ET PAR PÉRIODE
1987-2015 ET 2002-2015**

Variable Catégorie	Période 1987-2015		Période 2002-2015	
	LD + CD>20 % N=63 RIE (IC 95 %)	LD N=63 RIE (IC 95 %)	LD + CD>20 % N=40 RIE (IC 95 %)	LD N=40 RIE (IC 95 %)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	35 % [30-39]	-	37 % [31-42]	-
Longue durée (≥ 60 minutes)	4 % [3-5]	4 % [3-5]	4 % [3-5]	4 % [3-4]
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,54 (0,49;0,59)	-	0,54 (0,50;0,58)	-
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	5 % [4-6]	-	10 % [8-11]	-
Longue durée (≥ 60 minutes)	5 % [4-6]	5 % [4-6]	8 % [6-10]	8 % [6-9]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,04 (0,95;1,13)	1,04 (0,96;1,13)	1,14 (1,01;1,29)	1,16 (1,02;1,32)
Bourgo_comte	0,97 (0,89;1,05)	1,00 (0,92;1,08)	1,00 (0,91;1,09)	1,02 (0,93;1,12)
Bretagne	0,94 (0,83;1,07)	1,01 (0,88;1,15)	0,86 (0,71;1,03)	0,90 (0,76;1,08)
CALA	1,00 (0,89;1,11)	1,00 (0,88;1,12)	1,14 (1,01;1,27)	1,13 (1,00;1,28)
Centre	1,02 (0,92;1,12)	1,00 (0,91;1,10)	1,08 (0,95;1,24)	1,08 (0,95;1,22)
DOM Corse	0,87 (0,78;0,98)	0,92 (0,81;1,04)	0,98 (0,87;1,11)	1,02 (0,89;1,16)
IDF	1,00 (0,92;1,08)	1,01 (0,95;1,08)	1,03 (0,93;1,14)	1,05 (0,96;1,14)
MPLR	0,99 (0,90;1,07)	0,95 (0,87;1,05)	1,04 (0,95;1,14)	1,04 (0,94;1,14)
Nord	1,06 (0,95;1,18)	1,09 (0,96;1,23)	1,29 (1,07;1,54)	1,35 (1,12;1,63)
Normandie	1,15 (1,04;1,27)	1,14 (1,02;1,28)	1,10 (0,95;1,26)	1,13 (0,98;1,30)
P_Loire	0,90 (0,80;1,00)	0,88 (0,77;1,01)	0,90 (0,78;1,05)	0,88 (0,75;1,05)
PACA	1,25 (1,08;1,44)	1,18 (1,02;1,37)	1,05 (0,87;1,26)	0,98 (0,82;1,18)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Autre raison	0,98 (0,93;1,03)	0,97 (0,91;1,03)	0,86 (0,78;0,95)	0,84 (0,76;0,94)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Entreprise	0,91 (0,87;0,97)	0,97 (0,92;1,04)	0,89 (0,82;0,98)	0,92 (0,84;1,01)
Corps médical	1,02 (0,94;1,11)	1,04 (0,96;1,14)	0,99 (0,87;1,12)	0,93 (0,83;1,06)
Autre origine	0,88 (0,71;1,11)	0,89 (0,71;1,11)	0,81 (0,63;1,04)	0,85 (0,65;1,11)

Variable	Période 1987-2015		Période 2002-2015	
	LD + CD>20 % N=63 RIE (IC 95 %)	LD N=63 RIE (IC 95 %)	LD + CD>20 % N=40 RIE (IC 95 %)	LD N=40 RIE (IC 95 %)
Ventilation générale				
Non mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,07 (1,00;1,13)	1,07 (0,99;1,14)	0,98 (0,92;1,05)	0,97 (0,90;1,05)
Oui	1,14 (1,08;1,21)	1,07 (1,00;1,15)	1,08 (1,00;1,17)	0,99 (0,91;1,09)
Non	1,06 (1,01;1,12)	0,99 (0,94;1,05)	0,97 (0,90;1,04)	0,90 (0,82;0,98)
Information non relevée	0,99 (0,91;1,07)	0,95 (0,88;1,03)	0,88 (0,80;0,98)	0,85 (0,77;0,95)
Tranche d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,16 (1,07;1,26)	1,18 (1,08;1,29)	1,10 (1,00;1,22)	1,08 (0,97;1,20)
50-199 salariés	1,07 (0,99;1,17)	1,09 (1,00;1,19)	1,04 (0,94;1,16)	1,03 (0,92;1,15)
200 salariés ou plus	1,14 (1,02;1,27)	1,13 (1,00;1,28)	1,05 (0,91;1,21)	1,01 (0,87;1,17)
Information non relevée	1,10 (1,02;1,20)	1,14 (1,05;1,25)	1,05 (0,95;1,17)	1,11 (0,98;1,25)
Équipement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,67 (1,57;1,77)	1,72 (1,62;1,83)	1,86 (1,71;2,01)	1,88 (1,73;2,05)
Information non relevée	0,99 (0,90;1,09)	1,03 (0,94;1,12)	0,96 (0,86;1,08)	0,95 (0,86;1,06)
Type de procédé				
Ouvert		-	1,00 (Réf.)	1,00 (Réf.)
Semi-ouvert			0,72 (0,66;0,79)	0,74 (0,68;0,81)
Fermé			0,73 (0,66;0,81)	0,74 (0,66;0,83)
Autre type de procédé			0,80 (0,65;0,99)	0,84 (0,67;1,06)
Fréquence d'exposition				
Permanent			1,00 (Réf.)	1,00 (Réf.)
Fréquent			1,06 (1,00;1,12)	1,07 (1,01;1,13)
Occasionnel			0,88 (0,82;0,95)	0,81 (0,75;0,88)
Autre fréquence d'exposition			0,89 (0,74;1,07)	0,88 (0,71;1,08)

**ANNEXE V.8 : COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE)
PAR MÉTA ANALYSE DES AGENTS MESURÉS EN AMBIANCE SUR UNE
LONGUE DURÉE (LD) ASSOCIÉE AUX PRÉLÈVEMENTS COURTE DURÉE SI
LEUR PROPORTION EST SUPÉRIEURE À 20 % (CD>20 %) ET PAR PÉRIODE
1987-2015 ET 2002-2015**

Variable Catégorie	Période 1987-2015		Période 2002-2015	
	LD + CD>20 % N=59 RIE (IC 95 %)	LD N=59 RIE (IC 95 %)	LD + CD>20 % N=31 RIE (IC 95 %)	LD N=31 RIE (IC 95 %)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	30 % [17-41]	-	40 % [-72-79]	-
Longue durée (≥ 60 minutes)	5 % [4-6]	5 % [4-6]	5 % [4-6]	5 % [4-6]
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,53 (0,46;0,63)	-	0,49 (0,26;0,95)	-
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	5 % [3-8]	-	10 % [-3-21]	-
Longue durée (≥ 60 minutes)	4 % [3-5]	5 % [4-6]	8 % [6-9]	8 % [6-9]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,07 (0,97;1,19)	1,08 (0,97;1,20)	1,00 (0,90;1,12)	1,01 (0,90;1,13)
Bourgo_comte	1,00 (0,92;1,08)	1,00 (0,92;1,08)	0,98 (0,87;1,09)	0,98 (0,87;1,09)
Bretagne	0,92 (0,79;1,07)	0,94 (0,81;1,10)	0,99 (0,81;1,20)	0,99 (0,81;1,21)
CALA	1,08 (0,96;1,21)	1,09 (0,97;1,23)	1,11 (0,97;1,27)	1,11 (0,97;1,27)
Centre	1,06 (0,95;1,18)	1,07 (0,95;1,21)	1,11 (0,94;1,30)	1,10 (0,94;1,30)
DOM Corse	1,01 (0,87;1,16)	0,96 (0,82;1,13)	1,03 (0,86;1,24)	1,04 (0,86;1,25)
IDF	0,95 (0,86;1,06)	0,95 (0,85;1,06)	0,93 (0,84;1,03)	0,93 (0,84;1,03)
MPLR	1,03 (0,92;1,15)	1,02 (0,91;1,15)	1,09 (0,97;1,24)	1,10 (0,97;1,24)
Nord	1,10 (0,93;1,30)	1,08 (0,92;1,29)	1,09 (0,85;1,40)	1,07 (0,84;1,38)
Normandie	1,15 (1,04;1,27)	1,15 (1,04;1,28)	0,97 (0,85;1,11)	0,97 (0,85;1,11)
P_Loire	0,91 (0,79;1,04)	0,91 (0,79;1,05)	0,92 (0,78;1,08)	0,91 (0,77;1,07)
PACA	1,13 (1,02;1,25)	1,13 (1,02;1,25)	1,03 (0,85;1,25)	1,03 (0,85;1,25)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Autre raison	1,08 (1,01;1,17)	1,10 (1,02;1,19)	0,97 (0,87;1,09)	0,96 (0,86;1,08)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Entreprise	0,92 (0,86;0,98)	0,92 (0,86;0,98)	0,97 (0,89;1,05)	0,97 (0,89;1,06)
Corps médical	0,92 (0,85;1,00)	0,93 (0,85;1,01)	0,99 (0,88;1,12)	1,00 (0,88;1,13)
Autre origine	1,13 (0,92;1,37)	1,12 (0,92;1,36)	1,18 (0,70;2,02)	1,19 (0,70;2,02)

Variable	Période 1987-2015		Période 2002-2015	
	LD + CD>20 % N=59 RIE (IC 95 %)	LD N=59 RIE (IC 95 %)	LD + CD>20 % N=31 RIE (IC 95 %)	LD N=31 RIE (IC 95 %)
Ventilation générale				
Non, mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	1,20 (1,08;1,33)	1,22 (1,10;1,35)	1,09 (0,98;1,22)	1,08 (0,97;1,21)
Oui	1,18 (1,06;1,32)	1,19 (1,06;1,34)	1,08 (0,94;1,24)	1,06 (0,92;1,23)
Non	0,98 (0,93;1,04)	0,98 (0,92;1,03)	0,84 (0,76;0,94)	0,83 (0,74;0,93)
Information non relevée	1,00 (0,91;1,10)	0,99 (0,90;1,09)	0,89 (0,78;1,02)	0,90 (0,78;1,03)
Tranche d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,00 (0,89;1,11)	0,99 (0,89;1,12)	1,11 (0,99;1,23)	1,10 (0,99;1,23)
50-199 salariés	0,97 (0,89;1,05)	0,96 (0,88;1,04)	0,96 (0,87;1,07)	0,96 (0,86;1,06)
200 salariés ou plus	0,98 (0,88;1,09)	0,98 (0,87;1,10)	0,93 (0,84;1,02)	0,93 (0,85;1,02)
Information non relevée	1,03 (0,94;1,13)	1,04 (0,94;1,14)	1,02 (0,87;1,18)	1,01 (0,87;1,18)
Équipement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	1,62 (1,46;1,81)	1,64 (1,47;1,83)	1,39 (1,22;1,59)	1,37 (1,20;1,56)
Information non relevée	1,04 (0,96;1,12)	1,06 (0,98;1,15)	0,89 (0,80;1,00)	0,89 (0,79;0,99)
Type de procédé				
Ouvert	-	-	1,00 (Réf.)	1,00 (Réf.)
Semi-ouvert	-	-	0,96 (0,88;1,06)	0,96 (0,87;1,05)
Fermé	-	-	0,92 (0,81;1,04)	0,91 (0,80;1,03)
Autre type de procédé	-	-	0,89 (0,73;1,09)	0,88 (0,72;1,08)
Fréquence d'exposition				
Permanent	-	-	1,00 (Réf.)	1,00 (Réf.)
Fréquent	-	-	1,06 (0,99;1,13)	1,06 (0,99;1,13)
Occasionnel	-	-	0,96 (0,86;1,08)	0,95 (0,85;1,07)
Autre fréquence d'exposition	-	-	0,78 (0,66;0,93)	0,79 (0,66;0,94)

ANNEXE VI.1 POURCENTAGE DE DONNÉES INDIVIDUELLES INFÉRIEURES À LA LIMITE DE QUANTIFICATION, RÉPARTITION DE LA PROPORTION DE MESURES COURTE DURÉE (CD) ET LONGUE DURÉE (LD) PAR AGENT CHIMIQUE COMPTANT PLUS DE 1 000 RÉSULTATS SUR LA PÉRIODE 2007-2015 POUR COLCHIC, SCOLA ET LA COMPARAISON COLCHIC SCOLA

Agent chimique 2007 2015	COLCHIC				SCOLA				Comparaison COLCHIC SCOLA			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CT (%)	LT (%)
Poussières inhalables (A)	12107	3,1	4	96	10389	19	3	97	22386	11	3	97
Acétone (S)	7177	14	28	72	5399	30	55	45	12491	21	40	60
Fer (M)	6302	11	4	96	-	-	-	-	-	-	-	-
Styrène (A)	5150	8,6	28	72	-	-	-	-	-	-	-	-
Manganèse (M)	5034	14	4	96	-	-	-	-	-	-	-	-
Toluène (S)	5085	35	28	72	10146	50	38	62	15173	45	35	65
Bois poussières de (A)	4874	1,9	1	99	37385	22	3	97	42193	19	2	98
Xylène (S)	4205	35	31	69	12832	59	38	62	16988	53	36	64
Hydrocarbures aliphatiques et alicycliques (S)	4135	24	25	75	-	-	-	-	-	-	-	-
Formaldéhyde (A)	3986	4,3	28	72	-	-	-	-	-	-	-	-
Butanone (S)	4025	25	26	74	5415	34	53	47	9370	30	42	58
Chrome total (M)	3557	33	4	96	-	-	-	-	-	-	-	-
Nickel (M)	3454	33	4	96	-	-	-	-	-	-	-	-
Ethylbenzène (S)	3431	39	33	67	6493	66	34	66	9884	57	33	67
Cuivre (M)	3356	24	4	96	-	-	-	-	-	-	-	-
Acétate d'éthyle (S)	3401	24	25	75	-	-	-	-	-	-	-	-
Zinc (M)	3081	11	5	95	-	-	-	-	-	-	-	-
Plomb (M)	3017	47	4	96	9057	36	7	93	12065	39	6	94
Ethanol (S)	3001	24	25	75	-	-	-	-	-	-	-	-
Poussières alvéolaires (A)	2837	12	1	99	22383	28	2	98	25299	26	2	98

Agent chimique 2007 2015	COLCHIC				SCOLA				Comparaison COLCHIC SCOLA			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CT (%)	LT (%)
Quartz (A)	2614	35	1	99	21176	46	2	98	23774	45	2	98
Cristobalite (A)	2609	94	1	99	18719	92	2	98	21312	93	2	98
Acétate de n-butyle (S)	2353	26	32	68	-	-	-	-	-	-	-	-
Aluminium (M)	2222	30	4	96	-	-	-	-	-	-	-	-
Alcool isopropylique (S)	2221	23	30	70	-	-	-	-	-	-	-	-
Méthacrylate de méthyle (S)	1879	52	26	74	-	-	-	-	-	-	-	-
Cobalt (M)	1825	57	2	98	-	-	-	-	-	-	-	-
N-Hexane (S)	1642	45	27	73	3451	47	15	85	5059	46	19	81
1_Méthoxypropane-2-ol (S)	1468	33	22	78	2523	53	49	51	3982	46	39	61
Acétate de 2-méthoxy-1-méthyléthyle (S)	1378	31	26	74	2394	63	48	52	3759	51	40	60
4-Méthylpentane-2-one (S)	1352	36	35	65	1347	50	44	56	2699	43	39	61
Hydrocarbures aromatiques monocycliques (S)	1288	29	25	75	-	-	-	-	-	-	-	-
Chrome VI soluble (M)	1242	33	4	96	-	-	-	-	-	-	-	-
Chrome VI insoluble (M)	1161	27	4	96	-	-	-	-	-	-	-	-
Cadmium (M)	1148	75	2	98	-	-	-	-	-	-	-	-
Titane (M)	1089	14	6	94	-	-	-	-	-	-	-	-
Aérosol inhalable de fluides d'usinage (A)	1044	7,3	0	100	-	-	-	-	-	-	-	-
Dichlorométhane (S)	1064	40	20	80	1017	42	46	54	2041	41	33	67
Benzène (S)	1016	76	27	73	10507	79	9	91	11493	79	10	90
Tridymite (A)	-	-	-	-	4192	79	2	98	-	-	-	-
1,2,4-triméthylbenzène (S)	-	-	-	-	3298	63	38	62	-	-	-	-
Ammoniac anhydre (A)	-	-	-	-	3268	27	41	59	-	-	-	-
Méthanol (S)	-	-	-	-	2997	77	32	68	-	-	-	-
Acide chlorhydrique (A)	-	-	-	-	2611	69	95	4,6	-	-	-	-
N-heptane (S)	-	-	-	-	2603	56	44	56	-	-	-	-

Agent chimique 2007 2015	COLCHIC				SCOLA				Comparaison COLCHIC SCOLA			
	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CD (%)	LD (%)	n	<LoQ (%)	CT (%)	LT (%)
Phénol (A)	-	-	-	-	2291	64	44	56	-	-	-	-
Mésitylène (S)	-	-	-	-	2166	76	40	60	-	-	-	-
Cumène (S)	-	-	-	-	1933	92	49	51	-	-	-	-
Fluorure d'hydrogène (A)	-	-	-	-	1893	63	42	58	-	-	-	-
Chrome VI et ses composés (M)	-	-	-	-	1846	50	33	67	-	-	-	-
Fibres céramiques réfractaires (A)	-	-	-	-	1795	22	17	83	-	-	-	-
1,2,3-Triméthylbenzène (S)	-	-	-	-	1768	85	35	65	-	-	-	-
Sulfure d'hydrogène (A)	-	-	-	-	1660	94	39	61	-	-	-	-
Cyclohexane (S)	-	-	-	-	1580	52	26	74	-	-	-	-
Pentane (S)	-	-	-	-	1274	54	16	84	-	-	-	-
Chloroforme (S)	-	-	-	-	1242	43	26	74	-	-	-	-
Cyclohexanone (S)	-	-	-	-	1117	66	40	60	-	-	-	-

Famille des agents chimiques : (A) Autre famille, (M) Métal et (S) Solvant

ANNEXE VI.2 COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR MÉTA ANALYSE DE COLCHIC POUR LES FAMILLES D'AGENTS CHIMIQUES ET DES MODÈLES AVEC ET SANS VARIABLE ALÉATOIRE

Variable Catégorie	Aléatoire N=33 RIE (95 % CI)	Solvant N=16 RIE (95 % CI)	Métaux N=12 RIE (95 % CI)	Sans aléatoire N=33 RIE (95 % CI)
Durée du prélèvement				
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :				
Courte durée (< 60 minutes) ⁽¹⁾	34 % [27-41]	35 % [27-42]	-	26 % [16-35]
Longue durée (≥ 60 minutes)	5 % [3-6]	5 % [3-7]	4 % [2-6]	4 % [3-5]
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,42 (0,37;0,48)	0,41 (0,36;0,46)	-	0,46 (0,39;0,55)
Année du prélèvement				
Taux de réduction annuelle de l'exposition :				
Courte durée (< 60 minutes) ⁽¹⁾	7 % [3-11]	6 % [2-10]	-	9 % [4-13]
Longue durée (≥ 60 minutes)	7 % [5-9]	9 % [6-13]	4 % [1-7]	8 % [6-10]
Région				
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
ARA	1,19 (0,98;1,45)	1,12 (0,82;1,52)	1,90 (1,36;2,65)	1,32 (1,10;1,59)
Bourgo_comte	0,90 (0,75;1,06)	0,84 (0,66;1,07)	1,15 (0,86;1,54)	0,92 (0,77;1,11)
Bretagne	0,87 (0,72;1,05)	0,80 (0,65;0,99)	1,08 (0,79;1,48)	0,77 (0,63;0,94)
CALA	1,04 (0,90;1,21)	1,12 (0,92;1,36)	1,14 (0,84;1,53)	1,13 (0,94;1,36)
Centre	0,98 (0,85;1,14)	1,00 (0,84;1,20)	0,86 (0,66;1,11)	1,01 (0,83;1,22)
DOM Corse	0,96 (0,79;1,17)	1,05 (0,75;1,47)	0,97 (0,73;1,29)	0,96 (0,77;1,20)
IDF	1,04 (0,87;1,25)	0,94 (0,75;1,17)	1,27 (1,00;1,59)	1,03 (0,88;1,21)
MPLR	1,01 (0,85;1,21)	1,04 (0,80;1,35)	1,19 (0,92;1,54)	1,04 (0,86;1,27)
Nord	1,32 (1,02;1,72)	1,20 (0,88;1,63)	2,11 (1,27;3,48)	1,39 (1,04;1,84)
Normandie	1,15 (0,93;1,41)	1,22 (0,89;1,69)	1,64 (1,04;2,60)	1,29 (1,05;1,60)
P_Loire	0,86 (0,74;1,00)	0,92 (0,71;1,19)	0,90 (0,68;1,19)	0,85 (0,73;0,99)
PACA	0,92 (0,71;1,18)	0,84 (0,62;1,14)	1,10 (0,76;1,58)	0,95 (0,74;1,22)
Motif de la demande				
Risque d'exposition	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
Autre raison	0,86 (0,73;1,01)	0,84 (0,65;1,08)	0,68 (0,51;0,91)	0,79 (0,67;0,94)
Origine de la demande				
Organisme de prévention	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
Entreprise	0,97 (0,84;1,12)	0,86 (0,69;1,06)	0,87 (0,61;1,23)	0,96 (0,85;1,08)
Corps médical	0,98 (0,83;1,15)	1,11 (0,85;1,44)	0,77 (0,62;0,95)	0,82 (0,69;0,98)
Autre origine	0,94 (0,73;1,21)	0,98 (0,67;1,43)	0,67 (0,46;0,96)	0,78 (0,56;1,08)
Ventilation générale				
Non, mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
Oui et présence d'un captage localisé	1,02 (0,91;1,15)	1,03 (0,87;1,23)	0,80 (0,68;0,94)	1,02 (0,88;1,19)
Oui	1,08 (0,94;1,23)	1,15 (0,97;1,38)	1,14 (0,97;1,34)	1,22 (1,04;1,42)

Variable Catégorie	Aléatoire N=33 RIE (95 % CI)	Solvant N=16 RIE (95 % CI)	Métaux N=12 RIE (95 % CI)	Sans aléatoire N=33 RIE (95 % CI)
Non	0,95 (0,86;1,04)	0,95 (0,84;1,08)	1,06 (0,92;1,23)	1,02 (0,92;1,13)
Information non relevée	0,84 (0,77;0,92)	0,88 (0,77;1,00)	0,81 (0,68;0,97)	0,89 (0,76;1,04)
Tranches d'effectif				
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
20-49 salariés	1,10 (0,99;1,22)	1,04 (0,89;1,22)	1,17 (1,00;1,36)	1,02 (0,91;1,14)
50-199 salariés	1,01 (0,89;1,15)	1,01 (0,85;1,19)	1,16 (0,99;1,37)	1,05 (0,94;1,17)
200 salariés ou plus	1,08 (0,91;1,28)	1,09 (0,85;1,38)	1,09 (0,79;1,50)	1,08 (0,91;1,27)
Information non relevée	1,07 (0,94;1,22)	1,09 (0,87;1,38)	1,18 (0,93;1,52)	1,08 (0,95;1,22)
Équipement de protection individuelle				
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
Adapté	1,82 (1,63;2,03)	1,86 (1,56;2,20)	1,95 (1,64;2,32)	1,54 (1,34;1,76)
Information non relevée	1,04 (0,92;1,16)	1,05 (0,90;1,23)	0,78 (0,59;1,03)	0,99 (0,86;1,14)
Type de procédé				
Ouvert	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
Semi-ouvert	0,74 (0,65;0,84)	0,85 (0,70;1,03)	0,55 (0,44;0,70)	0,66 (0,57;0,77)
Fermé	0,71 (0,62;0,81)	0,77 (0,62;0,95)	0,64 (0,46;0,89)	0,66 (0,57;0,76)
Autre type de procédé	0,80 (0,59;1,08)	0,79 (0,60;1,02)	0,56 (0,29;1,06)	0,93 (0,63;1,35)
Fréquence d'exposition				
Permanent	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)	1,00 (1,00;1,00)
Fréquent	0,96 (0,89;1,04)	0,95 (0,83;1,08)	1,03 (0,92;1,16)	0,93 (0,87;0,99)
Occasionnel	0,80 (0,75;0,84)	0,82 (0,76;0,89)	0,62 (0,53;0,73)	0,79 (0,71;0,88)
Autre fréquence d'exposition	0,66 (0,52;0,85)	0,64 (0,47;0,88)	0,46 (0,25;0,84)	0,86 (0,58;1,26)

⁽¹⁾Données calculées avec 17, 16 et 17 agents chimiques respectivement pour Aléatoire, Solvant et Sans aléatoire

ANNEXE VI.3 COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR MÉTA ANALYSE DE SCOLA POUR LES SOLVANTS ET DES MODÈLES AVEC ET SANS VARIABLE ALÉATOIRE

Variable Catégorie	Aléatoire N=20 RIE (95 % CI)	Solvant N=12 RIE (95 % CI)	Sans aléatoire N=20 RIE (95 % CI)
Durée du prélèvement			
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :			
Courte durée (< 60 minutes) ⁽¹⁾	37 % [21-50]	35 % [14-50]	38 % [21-51]
Longue durée (≥ 60 minutes)	4 % [3-6]	4 % [3-6]	4 % [3-5]
Ratio Conc(240)/Conc(30)⁽¹⁾	0,39 (0,32;0,48)	0,40 (0,34;0,48)	0,39 (0,32;0,47)
Année du prélèvement			
Taux de réduction annuelle de l'exposition :			
Courte durée (< 60 minutes) ⁽¹⁾	-5 % [-14-3]	-8 % [-17-1]	-7 % [-16-2]
Longue durée (≥ 60 minutes)	0 % [-4-4]	2 % [-5-9]	0 % [-6-5]
Région			
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,09 (0,93;1,28)	1,05 (0,58;1,91)	1,11 (0,95;1,30)
Bourgo_comte	1,07 (0,89;1,28)	1,78 (1,10;2,89)	1,18 (0,96;1,46)
Bretagne	1,27 (1,00;1,61)	0,94 (0,51;1,72)	1,44 (1,07;1,94)
CALA	1,00 (0,84;1,19)	0,85 (0,47;1,52)	1,13 (0,88;1,46)
Centre	0,91 (0,73;1,15)	1,35 (0,63;2,91)	1,31 (0,98;1,74)
DOM Corse	1,22 (0,84;1,77)	1,13 (0,45;2,86)	1,26 (0,85;1,85)
IDF	1,04 (0,85;1,27)	0,94 (0,48;1,84)	1,17 (0,92;1,49)
MPLR	1,20 (1,02;1,42)	1,37 (0,82;2,28)	1,17 (1,00;1,37)
Nord	1,06 (0,88;1,29)	1,41 (0,82;2,43)	1,26 (0,97;1,64)
Normandie	0,99 (0,79;1,25)	1,11 (0,59;2,10)	1,15 (0,93;1,43)
P_Loire	1,18 (0,99;1,41)	1,49 (0,91;2,46)	1,31 (1,05;1,62)
PACA	1,12 (0,95;1,31)	1,31 (0,86;1,99)	1,09 (0,94;1,26)
Ventilation générale			
Non, mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	0,81 (0,72;0,91)	0,72 (0,61;0,85)	0,80 (0,72;0,89)
Oui	0,81 (0,70;0,93)	0,89 (0,71;1,12)	0,85 (0,74;0,98)
Non	0,80 (0,69;0,91)	0,74 (0,54;1,00)	0,83 (0,73;0,94)
Information non relevée	0,79 (0,70;0,89)	0,74 (0,56;0,97)	0,82 (0,72;0,93)

Variable Catégorie	Aléatoire N=20 RIE (95 % CI)	Solvant N=12 RIE (95 % CI)	Sans aléatoire N=20 RIE (95 % CI)
Tranches d'effectif			
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	0,88 (0,78;0,98)	0,81 (0,56;1,16)	0,82 (0,72;0,93)
50-199 salariés	0,85 (0,73;0,99)	0,84 (0,54;1,31)	0,76 (0,64;0,90)
200 salariés ou plus	0,65 (0,53;0,79)	0,53 (0,33;0,85)	0,61 (0,50;0,73)
Equipement de protection individuelle			
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	2,51 (2,05;3,06)	2,38 (1,94;2,91)	2,56 (2,02;3,24)
Information non relevée	1,03 (0,91;1,17)	0,89 (0,73;1,09)	1,03 (0,92;1,17)

⁽¹⁾Données calculées avec 12, 10 et 12 agents chimiques respectivement pour Aléatoire, Solvant et Sans aléatoire

ANNEXE VI.4 COMPARAISON DES INDICES D'EXPOSITION RELATIVE (RIE) PAR MÉTA ANALYSE POUR LA COMPARAISON DE COLCHIC ET SCOLA SANS INTERACTION ET AVEC INTERACTIONS POUR TOUS LES SECTEURS D'ACTIVITÉ ET RESTREINT AU SECTEURS COMMUNS AUX BDEP

Variable	Sans interaction	Tous secteurs d'activité	Secteurs d'activité communs
Catégorie	N=15 RIE (95 % CI)	N=15 RIE (95 % CI)	N=15 RIE (95 % CI)
Durée du prélèvement			
Taux de réduction de l'exposition provoquée par une augmentation de 30 minutes de la durée de prélèvement :			
Courte durée (< 60 minutes) ⁽¹⁾	40 % [45;34]	36 % [30;40]	34 % [29;40]
Longue durée (≥ 60 minutes)	4 % [3-5]	4 % [2-5]	4 % [3-5]
Ratio Conc(240)/Conc(30) en 2015⁽¹⁾	0,43 (0,40;0,47)	0,50 (0,47;0,53)	0,50 (0,46;0,55)
Année du prélèvement			
Taux de réduction annuelle de l'exposition :			
Courte durée (< 60 minutes) ⁽¹⁾	5 % [-1-9]	9 % [4-14]	11 % [5-16]
Longue durée (≥ 60 minutes)	7 % [4-10]	9 % [5-12]	8 % [4-11]
Source			
COLCHIC	1,00 (Réf.)	-	-
SCOLA	0,69 (0,58;0,82)	-	-
Région			
PCLA	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
ARA	1,08 (0,92;1,27)	0,90 (0,77;1,04)	0,84 (0,73;0,97)
Bourgo_comte	1,05 (0,89;1,23)	0,91 (0,78;1,05)	0,88 (0,76;1,03)
Bretagne	1,04 (0,88;1,23)	0,87 (0,72;1,05)	0,87 (0,72;1,06)
CALA	1,14 (0,93;1,39)	1,14 (0,91;1,43)	1,05 (0,88;1,25)
Centre	0,97 (0,78;1,20)	1,04 (0,82;1,33)	0,98 (0,77;1,24)
DOM Corse	1,02 (0,81;1,29)	0,94 (0,72;1,24)	0,89 (0,69;1,15)
IDF	1,01 (0,83;1,23)	0,91 (0,79;1,05)	0,92 (0,79;1,07)
MPLR	1,01 (0,87;1,18)	0,93 (0,77;1,11)	0,89 (0,77;1,02)
Nord	1,20 (0,95;1,52)	1,26 (0,83;1,91)	1,25 (0,82;1,90)
Normandie	1,05 (0,84;1,31)	0,97 (0,81;1,16)	0,97 (0,81;1,17)
P_Loire	1,20 (0,97;1,48)	0,86 (0,71;1,04)	0,83 (0,70;1,00)
PACA	1,10 (0,89;1,36)	0,84 (0,61;1,16)	0,83 (0,62;1,11)
Ventilation générale			
Non, mais présence d'un captage localisé	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Oui et présence d'un captage localisé	0,84 (0,77;0,91)	1,08 (0,95;1,22)	1,12 (0,98;1,29)
Oui	0,87 (0,77;0,97)	1,04 (0,95;1,13)	1,06 (0,96;1,16)
Non	0,83 (0,73;0,95)	0,94 (0,84;1,05)	0,96 (0,85;1,07)
Information non relevée	0,84 (0,76;0,93)	0,80 (0,68;0,95)	0,81 (0,69;0,96)

Variable	Sans interaction	Tous secteurs d'activité	Secteurs d'activité communs
Catégorie	N=15 RIE (95 % CI)	N=15 RIE (95 % CI)	N=15 RIE (95 % CI)
Tranches d'effectif			
1-19 salariés	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
20-49 salariés	1,01 (0,93;1,11)	1,06 (0,97;1,15)	1,06 (0,97;1,16)
50-199 salariés	0,94 (0,83;1,07)	0,91 (0,82;1,02)	0,90 (0,83;0,99)
200 salariés ou plus	0,82 (0,71;0,96)	0,98 (0,76;1,27)	0,96 (0,73;1,26)
Information non relevée	1,15 (0,92;1,45)	1,17 (0,95;1,43)	1,23 (0,99;1,53)
Equipement de protection individuelle			
Non adapté	1,00 (Réf.)	1,00 (Réf.)	1,00 (Réf.)
Adapté	2,40 (1,98;2,90)	1,94 (1,71;2,20)	1,96 (1,71;2,25)
Information non relevée	1,03 (0,93;1,14)	0,97 (0,86;1,10)	1,02 (0,88;1,19)

⁽¹⁾Données calculées avec 9 agents chimiques

ANNEXE VI.5 EVOLUTION DES RATIOS SCOLA/COLCHIC

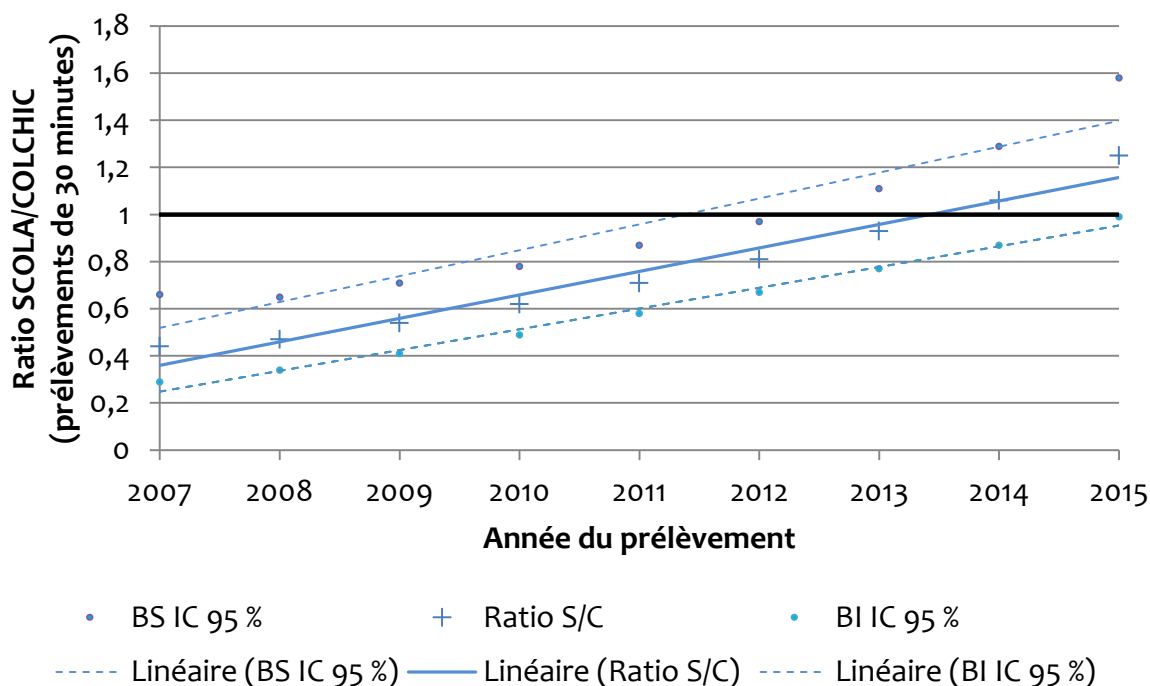


Figure AVI.3 Tendence temporelle du ratio SCOLA/COLCHIC pour les prélèvements d'une durée de 30 minutes avec les bornes inférieure et supérieure de l'intervalle de confiance à 95 %

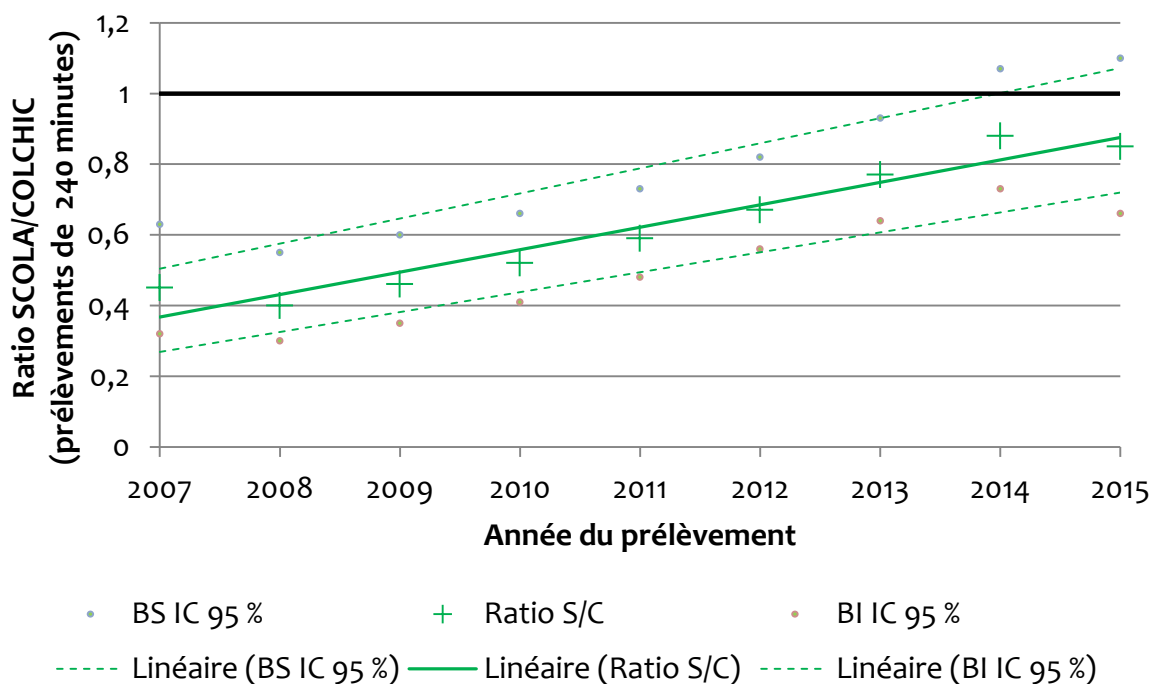


Figure AVI.4 Tendence temporelle du ratio SCOLA/COLCHIC pour les prélèvements d'une durée de 240 minutes avec les bornes inférieure et supérieure de l'intervalle de confiance à 95 %

Caractérisation des mesures d'exposition à des produits chimiques dans les bases de données françaises COLCHIC et SCOLA pour la prévention des maladies professionnelles

Résumé

En France, deux bases de données d'exposition professionnelle, COLCHIC et SCOLA, coexistent avec des objectifs différents (prévention et réglementation). Leur représentativité par rapport à la population générale est cependant inconnue, et fait l'objet de ce travail.

Après avoir effectué une analyse descriptive comparative, l'étude de l'association entre les niveaux d'exposition et les éléments contextuels a été réalisée par modélisation statistique pour chaque agent, séparément pour COLCHIC et SCOLA, puis dans un jeu de données commun. La synthèse à travers les agents s'est faite par Méta analyse.

COLCHIC et SCOLA contiennent respectivement 929 700 (670 agents chimiques) et 429 104 données (105). Trois forts prédicteurs « Durée de prélèvement », « Equipement de protection individuelle » et « Année » sont systématiquement associés aux niveaux dans les deux bases et 3 autres sont spécifiques à chacune d'elles. Avec des niveaux deux fois plus élevés dans COLCHIC comparés à SCOLA en 2007, leurs concentrations deviennent comparables entre 2012 et 2015.

COLCHIC et SCOLA représentent une source importante d'informations. La prise en compte des descripteurs associés aux mesures et l'utilisation de méthodes prédictives permettront d'en améliorer l'interprétation.

Mots clés :

COLCHIC, SCOLA, base de données d'exposition professionnelle, Modèles Tobit, Méta analyse, évaluation de l'exposition, hygiène du travail

Abstract

Two occupational exposure databases of occupational exposures to chemicals, COLCHIC and SCOLA, coexist in France with different objectives (prevention and compliance). Little is known about their representativeness of exposures in the general population. We explored to what extent COLCHIC and SCOLA adequately reflect occupational exposures in France.

After performing a descriptive and comparative analysis, associations between exposure levels and ancillary information were explored for each agent, separately for COLCHIC and SCOLA and in a common dataset, using statistical modelling. Modelling results were synthesized across agents using Meta analysis.

COLCHIC and SCOLA contain, respectively, 929 700 (670 chemicals) and 429 104 records (105). Three predictors "Sample Time", "Personal protective equipment" and "Year" are strongly associated with exposure levels across a large majority of chemicals in both databases, and 3 others are specific to each one. Exposure levels are in average twice higher in COLCHIC compared to SCOLA in 2007, but become comparable from 2012-2015.

COLCHIC and SCOLA are an important source of information. Inclusion of descriptors associated with exposure levels in our study and the use of predictive methods should help to improve their interpretation.

Keywords:

COLCHIC, SCOLA, occupational exposure database, Tobit models, Meta-analysis, exposure assessment, industrial hygiene