



Université de Lille - Ecole Gradué Biologie Santé de Lille

Champ disciplinaire : Recherche clinique, innovation technologique, santé publique

THESE POUR LE DIPLOME DE DOCTEUR EN SCIENCES

Etude épidémiologique des associations entre des expositions environnementales et les altérations métaboliques, en population générale

Présentée et soutenue publiquement le **16 janvier 2026 à 14 heures**
à l'Institut Pasteur de Lille

Par Marion HOSTENS (DARRAS-HOSTENS)

Jury

Président :

Monsieur le Professeur Philippe AMOUYEL

Rapporteurs :

Monsieur le Professeur Nicolas VENISSE

Madame le Docteur Nathalie BONVALLOT

Assesseurs :

Madame le Docteur Cécile CHEVRIER

Madame le Docteur Caroline LANIER

Directeur de thèse :

Monsieur le Docteur Luc DAUCHET

Liste des abréviations

4-HNE : 4-hydroxynonanal
8-OHdG : 8-hydroxydeoxyguanosine
ACP : Analyse en Composantes Principales
ADEME : Agence de l'Environnement et de la Maîtrise de l'Énergie
ADMS : Atmospheric Dispersion Modelling System
ANCOVA : Analyse de covariance
ANSES : Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail
ATMO Hauts-de-France : Organisme de surveillance de la qualité de l'air
CEA : Commissariat à l'Énergie Atomique et aux Énergies Alternatives
CEsilica : Cumulative Exposure Index to Silica (indice cumulatif d'exposition à la silice)
CC16 : Club cell secretory protein
CSM : Chromatographie couplée à la spectrométrie de masse
CO : Monoxyde de carbone
CRB : Centre de Ressources Biologiques
CRPus : Protéine C-réactive ultrasensible
CT : Cholestérol total
CVF : Capacité vitale fonctionnelle
ChEBI : Chemical Entities of Biological Interest
CRISTAL : Centre de Recherche en Informatique, Signal et Automatique de Lille
ECRIN : Environnement Climat Recherche & Innovation
EIS : Etude d'impact en santé
ELISABET : Enquête Littoral Souffle Air Biologie Environnement
ENNS : Etude nationale nutrition santé
ET : Écart-type
FDR : False Discovery Rate (taux de fausses découvertes)
FeNO : Fraction expirée du monoxyde d'azote
FOPs : Fluorescent oxidation products
GSEA : Gene Set Enrichment Analysis
GWAS : Genome-Wide Association Studies
HDL : High-Density Lipoprotein (Lipoprotéines de haute densité)
HMDB : Human Metabolome DataBase
HbA1c : Hémoglobine glyquée
IBE : indicateur biologique d'exposition
IC : Intervalle de Confiance
IMC : Indice de Masse Corporelle
INRAE : Institut National de la Recherche pour l'Agriculture, l'Alimentation et l'Environnement
INSEE : Institut National de la Statistique et des Études Économiques
INSERM : Institut National de la Santé et de la Recherche Médicale
IIQ : Intervalle interquartile
IRIS : Îlots Regroupés pour l'Information Statistique
JEM : Job Exposure Matrix (Matrice emplois-expositions)
KEGG : Kyoto Encyclopedia of Genes and Genomes
LASSO : Least Absolute Shrinkage and Selection Operator
MET : équivalent métabolique

LDL : Low-Density Lipoprotein (Lipoprotéines de basse densité)
MetaboAnalyst : Outil bioinformatique pour l'analyse métabolomique
Metalisabet : Projet de recherche sur les expositions environnementales et les perturbations métaboliques
MMLR : Multivariable multiple linear regression
MSEA : Metabolite Set Enrichment Analysis
NO : Monoxyde d'azote
NO₂ : Dioxyde d'azote
NO_x : Oxydes d'azotes
O₃ : Ozone
O-PLS-DA : Orthogonal Partial Least Squares-Discriminant Analysis
ORA : Over-Representation Analysis
PAD : Pression artérielle diastolique
PAS : Pression artérielle systolique
PFHpA : Perfluoroheptanoic acid
PFHxS : Perfluorohexane sulfonic acid
PFNA : Perfluorononanoic acid
PFUnDA : Perfluoroundecanoic acid
PFOS : Perfluorooctane sulfonate
PFOA : Perfluorooctanoic acid
PFAS : Per- and polyfluoroalkyl substances
PLS : Partial Least Squares (Moindres carrés partiels)
PLS-DA : Partial Least Squares – Discriminant Analysis
PM_{2,5} : Particules en suspension ayant un diamètre aérodynamique inférieur à 2,5 µm
PM₁₀ : Particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm
PREDIMED : PREvención con Dieta MEDiterránea
PROSPERO : Base internationale d'enregistrement des revues systématiques
RGPD : Règlement général sur la protection des données
RMN : Résonance Magnétique Nucléaire
ROC : Receiver operating characteristic
SAA : Sérum Amyloïde A
SM : Spectrométrie de masse
SO₂ : Dioxyde de soufre
sPLS : Sparse Partial Least Squares
ssPA : Single sample pathway analysis
TG : Triglycérides
THMs : Trihalométhanes
TVO : Troubles ventilatoires obstructifs
UMAP : Uniform manifold approximation and projection
VEMS : Volume Expiratoire Maximal par Seconde
VBR : Valeur de référence biologique
VLB : Valeur limite biologique
Z-score : Score standardisé

Remerciements

Au Président de jury,

Monsieur le Professeur Philippe Amouyel,

Professeur des Universités-Praticien Hospitalier
Epidémiologie, Economie de la Santé et Prévention
Université de Lille, INSERM, CHU Lille, Institut Pasteur de Lille, U1167

Vous me faites l'honneur de présider ce jury, acceptez pour cela mes plus sincères remerciements. Merci également pour votre accueil dans l'unité, votre confiance et votre soutien. C'est un plaisir de travailler dans votre équipe.

Aux rapporteurs,

Monsieur le Professeur Nicolas Venisse,

Professeur des Universités-Praticien Hospitalier UFR Santé
Laboratoire EBI, UMR CNRS 7267, responsable Équipe Interaction Homme Environnement Santé
Université de Poitiers
Service de Toxicologie et Pharmacocinétique
et CIC INSERM 1402, responsable axe EATHER/Santé Environnementale
CHU Poitiers

Toute ma gratitude pour votre participation à ce jury. Merci d'avoir accepté d'évaluer ce travail, pour la lecture attentive du manuscrit et la pertinence de vos commentaires.

Madame le Docteur Nathalie Bonvallot,

Enseignante chercheuse à l'EHESP
Directrice adjointe de l'IRSET, INSERM UMR_S1085
EHESP, département des sciences en santé environnementale

Toute ma reconnaissance pour votre participation à ce jury. Merci d'avoir accepté d'évaluer ce travail, de consacrer du temps à sa lecture et pour la richesse de vos remarques et suggestions.

Aux examinateurs,

Madame le Docteur Cécile Chevrier,

Co-responsable de l'équipe Elixir, DR Inserm
Épidémiologie et science de l'exposition en santé-environnement [Elixir]
IRSET

Pour votre participation aux CSI, vos précieux conseils et votre gentillesse.

Madame le Professeure Caroline Lanier,

Professeure des universités
CNU : SECTION 87 - SCIENCES BIOLOGIQUES, FONDAMENTALES ET CLINIQUES
Université de Lille

Pour avoir accepté de faire partie de ce jury. Recevez mes sincères remerciements.

A mon Directeur de thèse,

Monsieur le Docteur Luc Dauchet,

Maitre de conférences des Universités-Praticien Hospitalier, Université de Lille, CHU de Lille
Institut Pasteur de Lille, U1167

Merci pour ton aide dans cette thèse. Merci pour tes conseils, ta bonne humeur, ta patience, ta compréhension, ton soutien, les pièces de théâtre partagées ensemble. Tu n'es pas seulement mon directeur de thèse mais aussi un ami que j'apprécie beaucoup.

A Monsieur Aurélien Binet,

Pour votre participation à tous les CSI et vos précieux conseils qui m'ont permis d'aborder cette thèse avec sérénité.

A toute l'équipe du CEA,

En particulier François Fenaille et Florence Castelli pour votre aide tout au long de ce travail.

A toute l'équipe de l'INRAE de Toulouse

En particulier Fabien Jourdan et Clément Frainay pour votre accueil et la formation que vous m'avez dispensée. Merci aussi à Ludovic Cottret pour son aide et sa réactivité.

A mes collègues de l'Institut Pasteur de Lille (UMR1167-RID-AGE-Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement),

Pour votre bonne humeur et le travail que l'on fait ensemble.
Je remercie tout particulièrement Roberto et Delphine, vous m'avez beaucoup soutenue durant cette thèse. Merci aussi à Manon pour son aide pour le codage.

A tous les participants de l'enquête ELISABET

A l'équipe d'IMPECS, en particulier Jean-Marc Lo Guidice

A Jean DALLEGONVILLE, pour avoir initié le projet Metalisabet

A l'ADEME pour le financement du projet Metalisabet

A Nicolas, mon mari

Merci pour ton aide tout au long de cette thèse. Merci pour tout le bonheur que tu m'apportes et pour les beaux projets qu'on fait ensemble.
J'ai beaucoup de chance de t'avoir à mes côtés.

A mon père,

Merci pour ton soutien, et les moments de détente au restaurant passés ensemble.

A mes grands-mères,

Qui auraient aimé lire cette thèse.

A ma famille,

Pour votre présence et les moments que l'on partage.

A mes amies,

Avec qui je passe de bons moments.
Je pense notamment à Mathilde présente depuis toutes ces années. Malgré la distance tu es toujours là quand j'ai besoin de toi.
Mélanie, merci pour ton soutien, nos sorties et nos moments de rires.
Et aussi Céline, merci pour notre amitié et pour les bons moments partagés. Je sais que je peux toujours compter sur toi.

A Bob, Perceval et Léonor,

Pour votre humour

A l'association Laurette Fugain

A Stéphanie Fugain,

Merci d'avoir créé l'association. Merci pour ta gentillesse, ta générosité et toutes les valeurs que l'on partage. Merci pour la confiance que tu m'accordes.

A Flora,

Pour ta gentillesse et nos discussions sympathiques.

A Marie-Paule,

Pour ton accueil parmi les bénévoles et les événements que l'on fait ensemble.

Merci aux donateurs de moelle osseuse

Merci à tous ceux qui font preuve de générosité et de bienveillance

A Djemaa,

Infirmière extraordinaire, merci de faire vivre le souvenir de ma maman dans le service

A ma Maman,

Dont le souvenir m'accompagne chaque jour

Merci pour tout

Dziękuję Ci, Mamo, za wszystko

RESUME

Cette thèse explore les liens entre expositions environnementales (pollution de l'air, métaux lourds) et altérations métabolomiques en population générale, à partir de 3 175 habitants des zones urbaines de Lille et Dunkerque (étude ELISABET). Elle combine une approche épidémiologique classique et une analyse métabolomique non ciblée de métabolites urinaires mesurés par chromatographie liquide-spectrométrie de masse sur 1 000 non-fumeurs, pour étudier les effets biologiques précoces de l'environnement sur la santé. Les travaux préliminaires montrent une diminution du HDL avec l'exposition aux PM₁₀ et une augmentation de la CRP et du TNF α après exposition à court terme à la pollution de l'air, suggérant un processus inflammatoire de bas grade. Une revue systématique des méthodes statistiques en métabolomique environnementale révèle une forte hétérogénéité et l'absence de consensus, limitant la comparabilité des résultats, et propose des pistes pour adapter les analyses aux objectifs spécifiques. Les analyses des expositions montrent que les mesures indirectes, comme la pollution atmosphérique, compliquent la détection d'associations, tandis que le plomb est associé à de nombreuses voies métaboliques, même à faibles niveaux, sans valeur seuil détectée. Ces résultats confirment l'intérêt de la métabolomique pour détecter des effets biologiques infracliniques. Finalement, la thèse met en lumière la complexité de l'étude de l'interaction entre environnement et métabolisme, la nécessité d'une standardisation méthodologique et ouvre des perspectives pour identifier des médiateurs précoces des atteintes cardiovasculaires et respiratoires liées aux expositions environnementales.

Epidemiological study of associations between environmental exposures and metabolomic alterations in the general population

ABSTRACT

This thesis explores the links between environmental exposures (air pollution, heavy metals) and metabolomic alterations in the general population, based on 3,175 residents of the urban areas of Lille and Dunkirk (ELISABET study). It combines a classical epidemiological approach with untargeted metabolomic analysis of urinary metabolites measured by liquid chromatography–mass spectrometry in 1,000 non-smokers, to investigate the early biological effects of environmental exposures on health.

Preliminary findings show a decrease in HDL associated with PM₁₀ exposure and an increase in hsCRP and TNF α after short-term air pollution exposure, suggesting a low-grade inflammatory process. A systematic review of statistical methods in environmental metabolomics reveals high heterogeneity and lack of consensus, limiting comparability of results, and proposes ways to tailor analyses to specific objectives. Exposure analyses indicate that indirect measures, such as air pollution, complicate the detection of associations, whereas lead is associated with numerous metabolic pathways even at low levels, with no detectable threshold. These results confirm the value of metabolomics in detecting subclinical biological effects.

In conclusion, this thesis highlights the complexity of studying the interaction between environment and metabolism, underscores the need for methodological standardization, and opens perspectives for identifying early mediators of cardiovascular and respiratory impairments related to environmental exposures.

Mots-clés : Epidémiologie ; Métabolomique ; Environnement ; Santé humaine ; Etude en population

Keywords: Epidemiology; Metabolomics; Environment; Human Health; Population-Based Study

Laboratoire : U1167 - RID-AGE - Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement – Institut Pasteur de Lille – Equipe Public health and molecular epidemiology of cardiovascular diseases – 1 rue du Professeur Calmette 59000 Lille

Table des matières

I.	INTRODUCTION GENERALE	22
A.	Environnement Santé	22
1.	Définition(s) de l'environnement et impact sur la santé dans le Monde et en France.	22
a)	Définition(s) de l'environnement	22
b)	Impact de l'environnement sur la santé.....	23
2.	Contexte spécifique de l'enquête ELISABET	24
3.	Choix des expositions	25
a)	Pollution atmosphérique.....	25
b)	Exposition domestique.....	25
c)	Exposition professionnelle	26
d)	Exposition aux métaux lourds	27
4.	Etude des liens entre expositions environnementales et pathologies : rôle des intermédiaires	27
a)	Etude directe du lien entre exposition environnementale et santé	29
b)	Critères intermédiaires ciblés.....	29
(1)	Facteurs de risque connus.....	29
(2)	Biomarqueurs	30
(3)	Avantages et inconvénients des intermédiaires ciblés.....	30
c)	Critères intermédiaires non ciblés	31
B.	Apports de la métabolomique	31
1.	Définition de la métabolomique	31
2.	Méthodes de mesures.....	31
3.	Place de la bio-informatique	33
4.	Les avantages de la métabolomiques	33
5.	Les inconvénients de la métabolomique.....	33
6.	Les usages de la métabolomique	34
7.	Enjeux méthodologiques de l'analyse de données de métabolomique	35
C.	Objectif.....	36
II.	METHODE GENERALE	37
A.	Matériel	37
1.	Présentation de l'enquête ELISABET	37
a)	Recrutement des participants	37

b)	Mesures des expositions	38
c)	Pollution atmosphérique.....	38
d)	Exposition domestique	38
e)	Exposition professionnelle	38
f)	Expositions aux métaux lourds.....	39
2.	Présentation de Metalisabet	39
B.	Méthode de la thèse : stratégie des analyses	40
1.	Etudier les relations entre exposition et pathologie à l'aide d'un intermédiaire concernant les facteurs de risque cardiovasculaire	40
2.	Etudier ces associations avec un intermédiaire concernant des biomarqueurs.....	41
3.	Revue méthodologique métabolomique	41
4.	Analyses métabolomiques	41
5.	Analyses développées sur l'exposition au plomb.....	43
III.	LOTS DE TRAVAIL (work packages)	44
A.	Etudier les relations entre exposition et pathologie à l'aide d'un intermédiaire concernant les facteurs de risque cardiovasculaire.	44
1.	Introduction.....	44
2.	Méthode.....	46
a)	Population	46
b)	Données sur les polluants	46
c)	Données socio-économiques sur les quartiers	46
d)	Biomarqueurs lipidiques et pression artérielle	46
e)	Analyses statistiques	47
3.	Résultats de cette étude.....	48
a)	Population	48
b)	Pollution	49
c)	Associations entre pollution de l'air résidentielle et facteurs de risque cardiovasculaire.....	49
d)	Analyses de sensibilité.....	49
4.	Discussion de cette étude	54
5.	Conclusion de cette étude.....	55
B.	Etudier les relations entre exposition et pathologie à l'aide d'un intermédiaire de type biomarqueur.....	56
1.	Introduction.....	56
2.	Méthode.....	58
a)	Population	58

b)	Zones d'étude	59
c)	Pollution atmosphérique et mesures météorologiques	59
d)	Biomarqueurs	60
e)	Analyses statistiques	61
3.	Résultats de cette étude.....	62
a)	Population	62
b)	Pollution	63
c)	Associations entre la pollution et les biomarqueurs inflammatoires	63
(1)	Association entre la pollution et la CRPus :.....	63
(2)	Association entre la pollution et les cytokines :.....	64
d)	Analyses de sensibilité.....	64
4.	Discussion de cette étude	73
a)	CRPus.....	73
b)	Cytokines	76
c)	Conclusion de cette étude.....	78
C.	Revue de la littérature :.....	80
1.	Introduction.....	80
2.	Méthode.....	80
3.	Résultats de cette étude.....	83
a)	Sélection des publications	83
b)	Questions scientifiques	83
c)	Hypothèses scientifiques.....	84
d)	Exposition et critère de jugement	85
e)	Analyses statistiques	85
(1)	Méthodes d'analyses unidimensionnelles	85
(2)	Méthodes d'analyses multidimensionnelles	85
f)	Représentations graphiques.....	85
g)	Analyses statistiques selon la question scientifique	86
(1)	Question scientifique relative aux processus ou voies métaboliques	86
(2)	Question scientifique relative à la médiation.....	86
h)	Relation entre la question scientifique et l'hypothèse H1, ainsi qu'entre l'hypothèse H1 et le critère de jugement statistique	86
(1)	Relation entre la question scientifique et l'hypothèse H1	86
(2)	Relation entre l'hypothèse H1 et l'outcome statistique.....	87
i)	Validité statistique	88
4.	Discussion de cette étude	122

a)	Résultats principaux	122
b)	Représentations graphiques.....	122
c)	La relation entre la question scientifique et le type d'analyse.....	123
(1)	Question scientifique : « Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? » - analyses unidimensionnelles.	123
(2)	Question scientifique : « Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ? » – calcul de la proportion médiée	127
(3)	Analyses multidimensionnelles – question scientifique : « déterminer si l'exposition altère le métabolisme global ».....	128
d)	Forces et limites de cette revue	130
5.	Conclusion	131
D.	Associations entre expositions environnementales et métabolomique.....	132
1.	Introduction.....	132
2.	Méthodes	132
3.	Résultats	134
a)	Population	134
b)	Variables de démonstrations et expositions	135
c)	Métabolites	136
d)	Analyses.....	136
(1)	Preuve de concept : variables de démonstration.....	136
(2)	Pollution atmosphérique.....	141
(3)	Exposition domestique.....	166
(4)	Exposition professionnelle	171
(5)	Métaux lourds : cadmium	191
4.	Discussion de ces résultats.....	196
E.	Associations entre plombémie et métabolites et voies métaboliques.....	198
1.	Introduction.....	198
a)	Exposition au plomb.....	198
(1)	Les organes touchés par la toxicité du plomb :.....	198
(2)	Les valeurs de références :	199
b)	Littérature.....	200
c)	Objectif.....	201
2.	Méthode.....	201
a)	Population, mesure de l'exposition, mesures des métabolites.....	201
b)	Analyses statistiques	201

3.	Résultats de cette étude.....	202
a)	Population	202
b)	Exposition au plomb.....	202
c)	Variables métaboliques	203
d)	Analyses par métabolites individuels.....	203
e)	Analyses par voies métaboliques	206
(1)	ORA (MetExplore).....	206
(2)	MSEA	207
(3)	Z-score	207
(4)	Rapport de vraisemblance	208
(5)	Comparaisons des voies métaboliques selon les méthodes.....	209
f)	Relation effet-dose	211
4.	Discussion de cette étude	212
a)	Comparaison des 4 méthodes pour étudier les voies métaboliques.....	212
b)	Comparaison avec la littérature	213
c)	Physiopathologie	213
d)	Relation effet-dose	215
5.	Conclusion de cette étude.....	216
IV.	DISCUSSION GENERALE	217
A.	Etudes de variables intermédiaires : facteurs de risque connus et biomarqueurs....	217
1.	Etude d'intermédiaires : facteurs de risque.....	217
2.	Etude d'intermédiaires : biomarqueurs	217
B.	Métabolomique.....	218
1.	Fishing expedition	218
2.	Importance de poser les hypothèses	218
C.	Etudes des liens entre expositions et métabolomique dans ELISABET.....	221
1.	Choix des analyses sur l'ensemble des expositions.....	221
2.	Discussion des résultats obtenus	221
3.	Exposome	222
4.	Exemple du plomb.....	222
D.	Perspectives.....	222
V.	CONCLUSION	225
VI.	REFERENCES	226
VII.	ANNEXES.....	255

Table des tableaux

Tableau III.A.3.1. Caractéristiques des participants de l'étude	p. 50
Tableau III.A.3.2. Associations entre l'exposition aux polluants atmosphériques et les biomarqueurs sériques (g/L)	p. 51
Tableau III.A.3.3. Associations entre l'exposition aux polluants atmosphériques et la pression artérielle (mmHg)	p. 52
Tableau III.A.3.4. Associations entre l'exposition aux polluants atmosphériques et les biomarqueurs sériques (g/L) – analyses de sensibilité	p. 53
Tableau III.B.3.1. Caractéristiques des participants de l'étude	p. 66
Tableau III.B.3.2. Associations entre l'exposition aux polluants atmosphériques et le taux sérique de CRPus (mg/L)	p. 67
Tableau III.B.3.3. Associations entre l'exposition à la pollution atmosphérique et les taux plasmatiques de biomarqueurs (pg/mL)	p. 68
Tableau supplémentaire III.B.3.1. Caractéristiques des participants, selon la classe d'IMC (kg/m ²) (N = 3074)	p. 69
Tableau supplémentaire III.B.3.2. Coefficients de Pearson pour les corrélations entre les taux de polluants atmosphériques résidentiels (µg/m ³) en fonction des différentes années à Lille	p. 69
Tableau supplémentaire III.B.3.3. Associations entre l'exposition à la pollution atmosphérique à court terme et le taux sérique de CRPus (mg/L) – modèles multipolluants	p. 70
Tableau supplémentaire III.B.3.4. Associations entre l'exposition à la pollution atmosphérique résidentielle et le taux sérique de CRPus (mg/L) – modèles multipolluants	p. 71
Tableau III.C.3.1. Description des études	p. 89
Tableau III.C.3.2. Évaluation de la qualité des études et du risque de biais (score de ROBINS)	p. 93
Tableau III.C.3.3. Cohérence entre la question scientifique, l'hypothèse H1, l'outcome statistique et l'absence de problème statistique	p. 95
Tableau supplémentaire III.C.3.1. Descriptions détaillées des populations d'étude	p. 101
Tableau supplémentaire III.C.3.2. Cohérence entre la question scientifique, H1 et l'outcome statistique	p. 105
Tableau supplémentaire III.C.3.3. Description des expositions et des mesures métabolomiques	p. 113
Tableau III.D.3.1. Caractéristiques des participants de l'étude	p. 132
Tableau III.D.4.1 : Synthèse	p.194
Tableau III.E.3.1. Caractéristiques des participants de l'étude	p. 199
Tableau III.E.3.2. Description de la concentration sanguine en plomb selon le sexe	p. 200
Tableau III.E.3.3. Métabolites associés à la concentration sanguine au plomb (µg/L)	p. 201
Tableau III.E.3.4. Comparaisons des voies métaboliques associées à la plombémie selon les méthodes statistiques	p. 207
Tableau IV.B.2.1 : Stratégie d'analyse en métabolomique	p. 217

Table des figures

Figure I.A.1.1 : Définitions de l'environnement.....	p. 22
Figure I.A.4.1 : Etudes des liens entre expositions environnementales et différentes variables dans l'enquête ELISABET	p. 28
Figure II.B.4.1 : Schéma du workflow sur l'approche métabolomique	p. 42
Figure III.B.3.1 : Carte du Nord de la France	p. 72
Figure III.C.3.1 : Diagramme de flux (Flow chart)	p. 117
Figure III.C.3.2 : Cadre proposé pour adapter l'analyse statistique à l'objectif de l'étude p.	118
Figure III.D.3.1 : Volcano plot représentant les associations entre la consommation de fruits et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 133
Figure III.D.3.2 : QQ plot représentant les associations entre la consommation de fruits et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 134
Figure III.D.3.3 : Volcano plot représentant les associations entre la consommation de fruits et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p.135
Figure III.D.3.4 : QQ plot représentant les associations entre la consommation de fruits et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine p.	136
Figure III.D.3.5 : QQ plot représentant les p des rapports de vraisemblance des associations entre la consommation de fruits et les voies métaboliques (Z-score) ajustées sur l'âge, le sexe et la communauté urbaine	p. 137
Figure III.D.3.6 : Volcano plot représentant les associations entre l'exposition à court terme aux PM10 et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation p.	138
Figure III.D.3.7 : QQ plot représentant les associations entre l'exposition à court terme aux PM10 et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation p.	139
Figure III.D.3.8 : Volcano plot représentant les associations entre l'exposition à court terme aux PM10 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation p.	140
Figure III.D.3.9 : QQ plot représentant les associations entre l'exposition à court terme aux PM10 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation p.	141
Figure III.D.3.10 : QQ plot représentant les p des rapports de vraisemblance des associations entre l'exposition à court terme aux PM10 et les voies métaboliques (Z score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation.....	p. 142

Figure III.D.3.11 : Volcano plot représentant les associations entre l'exposition résidentielle aux PM10 et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 143
Figure III.D.3.12 : QQ plot représentant les associations entre l'exposition résidentielle aux PM10 et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 144
Figure III.D.3.13 : Volcano plot représentant les associations entre l'exposition résidentielle aux PM10 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 145
Figure III.D.3.14 : QQ plot représentant les associations entre l'exposition résidentielle aux PM10 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 146
Figure III.D.3.15 : QQ plot représentant les p des rapports de vraisemblance des associations entre l'exposition résidentielle aux PM10 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 147
Figure III.D.3.16 : Volcano plot représentant les associations entre l'exposition à court terme au NO2 et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation	p. 148
Figure III.D.3.17 : QQ plot représentant les associations entre l'exposition à court terme au NO2 et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation	p. 149
Figure III.D.3.18 : Volcano plot représentant les associations entre l'exposition à court terme au NO2 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation	p. 150
Figure III.D.3.19 : Volcano plot représentant les associations entre l'exposition à court terme au NO2 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation	p. 151
Figure III.D.3.20 : QQ plot représentant les p des rapports de vraisemblance des associations entre l'exposition à court terme au NO2 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation.....	p. 152
Figure III.D.3.21 : Volcano plot représentant les associations entre l'exposition résidentielle au NO2 et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 153

Figure III.D.3.22 : QQ plot représentant les associations entre l'exposition résidentielle au NO ₂ et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 154
Figure III.D.3.23 : Volcano plot représentant les associations entre l'exposition résidentielle au NO ₂ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 155
Figure III.D.3.24 : QQ plot représentant les associations entre l'exposition résidentielle au NO ₂ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation.....	p. 156
Figure III.D.3.25 : QQ plot représentant les p des rapports de vraisemblance des associations entre l'exposition résidentielle au NO ₂ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation	p. 157
Figure III.D.3.26 : Volcano plot représentant les associations entre l'exposition à court terme à l'O ₃ et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation	p. 158
Figure III.D.3.27 : QQ plot représentant les associations entre l'exposition à court terme à l'O ₃ et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation.....	p. 159
Figure III.D.3.28 : Volcano plot représentant les associations entre l'exposition à court terme à l'O ₃ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation	p. 160
Figure III.D.3.29 : QQ plot représentant les associations entre l'exposition à court terme à l'O ₃ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation	p. 161
Figure III.D.3.30 : QQ plot représentant les p des rapports de vraisemblance des associations entre l'exposition à court terme à l'O ₃ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation.....	p. 162
Figure III.D.3.31 : Volcano plot représentant les associations entre les moisissures et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 163
Figure III.D.3.32 : QQ plot représentant les associations entre les moisissures et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 164
Figure III.D.3.33 : Volcano plot représentant les associations entre les moisissures et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 165
Figure III.D.3.34 : QQ plot représentant les associations entre les moisissures et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 166

Figure III.D.3.35 : QQ plot représentant les p des rapports de vraisemblance des associations entre les moisissures et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 167
Figure III.D.3.36 : Volcano plot représentant les associations entre les vapeurs et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 169
Figure III.D.3.37 : QQ plot représentant les associations entre les vapeurs et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 170
Figure III.D.3.38 : Volcano plot représentant les associations entre les vapeurs et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 171
Figure III.D.3.39 : QQ plot représentant les associations entre les vapeurs et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 172
Figure III.D.3.40 : QQ plot représentant les p des rapports de vraisemblance des associations entre les vapeurs et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 173
Figure III.D.3.41 : Volcano plot représentant les associations entre les poussières et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 174
Figure III.D.3.42 : QQ plot représentant les associations entre les poussières et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 175
Figure III.D.3.43 : Volcano plot représentant les associations entre les poussières et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 176
Figure III.D.3.44 : QQ plot représentant les associations entre les poussières et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 177
Figure III.D.3.45 : QQ plot représentant les p des rapports de vraisemblance des associations entre les poussières et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 178
Figure III.D.3.46 : Volcano plot représentant les associations entre les gaz et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 179
Figure III.D.3.47 : QQ plot représentant les associations entre les gaz et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 180
Figure III.D.3.48 : Volcano plot représentant les associations entre les gaz et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 181
Figure III.D.3.49 : QQ plot représentant les associations entre les gaz et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 182
Figure III.D.3.50 : QQ plot représentant les p des rapports de vraisemblance des associations entre les gaz et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 183
Figure III.D.3.51 : Volcano plot représentant les associations entre les fumées et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine	p. 184
Figure III.D.3.52 : QQ plot représentant les associations entre les fumées et les métabolites, ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 185
Figure III.D.3.53 : Volcano plot représentant les associations entre les fumées et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 186
Figure III.D.3.54 : QQ plot représentant les associations entre les fumées et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine	p. 187
Figure III.D.3.55 : QQ plot représentant les p des rapports de vraisemblance des associations entre les fumées et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine.....	p. 188

Figure III.D.3.56 : Volcano plot représentant les associations entre la concentration sanguine en cadmium et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 189
Figure III.D.3.57 : QQ plot représentant les associations entre la concentration sanguine en cadmium et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 190
Figure III.D.3.58 : Volcano plot représentant les associations entre la concentration sanguine en cadmium et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 191
Figure III.D.3.59 : QQ plot représentant les associations entre la concentration sanguine en cadmium et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 192
Figure III.D.3.60 : QQ plot représentant les p des rapports de vraisemblance des associations entre la concentration sanguine en cadmium et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 193
Figure III.E.3.1 : Volcano plot représentant les associations entre la concentration sanguine en plomb et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 202
Figure III.E.3.2 : QQ plot représentant les associations entre la concentration sanguine en plomb et les métabolites, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 203
Figure III.E.3.3 : Volcano plot représentant les associations entre la concentration sanguine en plomb et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 204
Figure III.E.3.4 : QQ plot représentant les associations entre la concentration sanguine en plomb et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse.....	p. 205
Figure III.E.3.5 : QQ plot représentant les p des rapports de vraisemblance des associations entre la concentration sanguine en plomb et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse	p. 206
Figure III.E.3.6 : Courbe de la relation effet-dose entre la plombémie et la voie métabolique du glutathion	p. 208
Figure III.E.3.7 : Courbe de la relation effet-dose entre la plombémie (logarithmique) et la voie métabolique du glutathion	p. 209

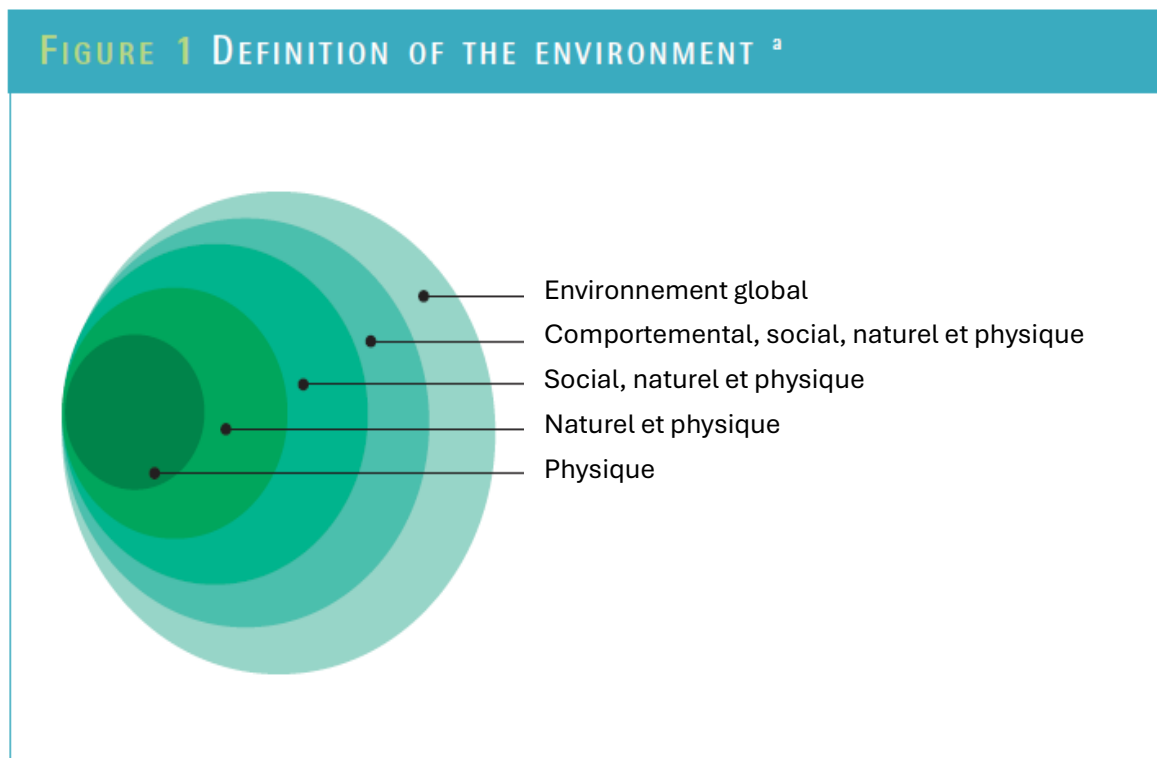
I. INTRODUCTION GENERALE

A. Environnement Santé

1. Définition(s) de l'environnement et impact sur la santé dans le Monde et en France.

a) Définition(s) de l'environnement

L'environnement peut être défini de différentes manières, allant de notions très larges jusqu'à des notions plus spécifiques. Ainsi, l'environnement peut être défini par l'environnement physique. Il peut aussi inclure l'environnement naturel et physique ; ou encore l'environnement social, physique et naturel. La définition peut être encore plus englobante en incluant l'environnement comportemental, social, naturel et physique (1,2). Ces définitions sont schématisées dans la figure I.A.1.1 ci-dessous.



^a (Adapted from Smith, Corvalán and Kjellström, 1999)

(2)

Figure I.A.1.1 : Définitions de l'environnement

Pour cette thèse, une définition large de l'environnement, intégrant les environnements comportementaux, sociaux, naturels et physiques, sera retenue, notamment pour les questions méthodologiques qui font l'objet d'un travail de revue systématique. Concernant les analyses originales, les facteurs environnementaux étudiés spécifiquement sont ceux pour lesquels a été conçue l'enquête ELISABET en population générale dont sont issues les données qui ont permis ce travail de thèse.

Il existe des méthodes d'estimation de la fraction attribuable de la morbi-mortalité liée aux causes environnementales.

b) Impact de l'environnement sur la santé.

La quantification de l'impact de l'environnement sur la santé (burden) est complexe. Les estimations sont très dépendantes des définitions et des données disponibles. Cette quantification peut être très variable selon les sources et reste souvent limitée à des ordres de grandeur.

Cette quantification repose sur la fraction attribuable en population qui est la proportion de cas ou de décès d'une maladie, dans la population générale, que l'on peut attribuer à un facteur de risque ou à une exposition. L'estimation de la fraction attribuable de la morbi-mortalité liée à l'environnement est un enjeu essentiel pour la prévention et le pilotage des politiques publiques. Les estimations réalisées dépendent de nombreux facteurs comprenant, entre autres, le choix des expositions retenues et leur mesure, le choix de la catégorie d'exposition la plus faible ou les données de la littérature scientifique, les modèles utilisés, etc. Chacune des estimations doit être interprétée en fonction de son contexte. La part attribuable de mortalité et de morbidité liée aux expositions environnementales nécessite le recours à des méthodes d'étude d'impact en santé (EIS). Ces méthodes tiennent compte des expositions de la population, des indicateurs sanitaires (incidence, prévalence, mortalité), des associations connues entre les expositions et les risques de pathologies (par exemple le risque relatif dans des études de cohorte).

Ces estimations nécessitent de faire l'hypothèse d'un lien causal entre l'exposition environnementale et la maladie. De plus, elles sont dépendantes des données et hypothèses retenues pour le calcul comme la prévalence de l'exposition ou la force de l'association estimée d'après la littérature. Pour ces raisons, les parts de risque attribuable à chaque exposition environnementale peuvent varier entre les sources (3). Les estimations des impacts environnementaux données dans la suite de ce travail doivent être interprétées en tenant compte de ces limites.

Le dernier rapport de l'Organisation Mondiale de la Santé (OMS) sur l'évaluation mondiale de la charge liée aux risques environnementaux (2016) a montré que 23 % des décès mondiaux (et 26 % des décès chez les enfants de moins de cinq ans) étaient dus à des facteurs environnementaux modifiables (4).

Les facteurs inclus étaient les suivants : (4)

- La pollution de l'air (y compris la fumée du tabagisme passif), de l'eau ou du sol par des agents chimiques ou biologiques.
- Les rayonnements ultraviolets (en particulier, la protection contre ceux-ci) et ionisants
- Le bruit, les champs électromagnétiques.
- Les risques professionnels, y compris les risques physiques, chimiques, biologiques et psychosociaux, ainsi que les conditions de travail.
- Les environnements bâtis, y compris le logement, les lieux de travail, l'aménagement du territoire et les routes.

- Les méthodes agricoles.
- Les changements climatiques et écosystémiques d'origine humaine.
- Les comportements liés aux facteurs environnementaux, par exemple la disponibilité d'une eau potable pour se laver les mains, ou encore l'activité physique favorisée par une meilleure conception urbaine.

Les facteurs non inclus étaient les suivants : (4)

- Consommation d'alcool et de tabac.
- Alimentation (sauf lorsqu'elle est liée à la dégradation de l'environnement).
- Milieux naturels des vecteurs qui ne peuvent raisonnablement être modifiés (par exemple, zones humides, lacs).
- Moustiquaires imprégnées d'insecticide (dans cette étude, elles sont considérées comme des interventions non environnementales).
- Chômage (dans la mesure où il n'est pas lié à la dégradation de l'environnement, aux maladies professionnelles, etc.).
- Agents biologiques naturels, tels que le pollen.
- Transmission interhumaine qui ne peut raisonnablement être évitée par des interventions environnementales, comme l'amélioration du logement, l'introduction de mesures d'hygiène sanitaire ou l'amélioration des conditions de travail.

2. Contexte spécifique de l'enquête ELISABET

Cette thèse s'appuie sur l'enquête ELISABET (Enquête Littoral Souffle Air Biologie Environnement) qui est une étude transversale portant sur un échantillon représentatif d'habitants de 40 à 65 ans des communautés urbaines de Lille et de Dunkerque, recrutés entre 2011 et 2013. Cette étude avait pour objectif principal de comparer la prévalence des troubles ventilatoires obstructifs (TVO) entre les communautés urbaines de Lille et de Dunkerque. L'enquête ELISABET est détaillée dans la section Méthode générale.

La zone de Lille est une zone urbaine dense, comportant seulement quelques espaces non urbains entre les trois principales villes (Lille, Roubaix et Tourcoing) et deux plus petites villes (Armentières et Comines). Le réseau routier et autoroutier y est très dense. La zone urbaine de Lille est traversée par cinq autoroutes ; en 2012, le nombre moyen de véhicules par jour variait de 36 000 à 176 000 (5,6).

La zone de Dunkerque est une zone urbaine côtière. La mer se trouve au nord de la ville, tandis qu'un port et une zone industrielle sont situés au nord-ouest. Les vents dominants soufflent du sud-ouest ; par conséquent, la pollution atmosphérique provenant de la zone industrielle est en grande partie poussée vers la mer. Le réseau routier et autoroutier de cette zone est également dense (7). La zone urbaine de Dunkerque est traversée par deux autoroutes ; en 2012, le nombre moyen de véhicules par jour variait de 27 000 à 61 000 (5,6).

Au-delà de la comparaison entre Lille et Dunkerque cette étude a vocation à s'intéresser aux expositions environnementales mesurables et ayant un poids important sur la santé en population générale.

Elle s'intéresse en particulier à la pollution atmosphérique. Elle exclut des expositions concernant des groupes spécifiques, comme des expositions professionnelles concernant des catégories ciblées (par exemple, les pesticides chez les agriculteurs, l'exposition à l'amiante) ou des expositions qui sont difficiles à mesurer dans cette population comme les perturbateurs endocriniens.

Les choix des expositions étudiées sont décrits ci-après. Une exposition peut se définir par le contact ou la mise en relation d'un individu ou d'une population avec un facteur susceptible d'avoir un effet sur la santé.

3. Choix des expositions

a) Pollution atmosphérique

Impact dans le Monde :

La pollution de l'air (extérieur et intérieur) est responsable, d'après l'Organisation Mondiale de la Santé, de 7 millions de décès prématurés chaque année (8). Cette estimation comprend les polluants suivants : les particules fines (PM), les oxydes d'azotes (Nox) dont le dioxyde d'azote (NO₂), l'ozone (O₃), le monoxyde de carbone (CO) et le dioxyde de soufre (SO₂).

Impact en France :

En France, en 2016, la part de décès attribuée aux particules en suspension (ou particules fines) était de 9 % correspondant à 48 000 décès chaque année ou encore à 950 000 années de vie perdues par an (9). Sur la période de 2016 à 2019, une estimation de 40 000 décès annuels en lien avec les PM_{2,5} était faite en France (10).

La pollution atmosphérique a des effets délétères sur de nombreux organes, entraînant ainsi des maladies pulmonaires (11–15), cardiovasculaires (11,16), cérébrales et psychiques (11,17), des cancers (18–20), ou encore des troubles métaboliques (21–25) ou de la reproduction (11).

La pollution atmosphérique fait partie des 3 premières causes en termes de causes environnementales de décès en France, avec le tabac et l'alcool. En effet, le nombre de décès attribuables au tabagisme était de 75 000, soit 13 % des décès en France en 2016, d'après le bulletin épidémiologique hebdomadaire de 2025 (26).

Le nombre de décès attribuables à l'alcool était de 41 000, soit 7% des décès en France en 2015 (27).

Dans ce travail, nous nous sommes intéressés aux polluants atmosphériques pour lesquels des données d'ATMO Hauts-de-France étaient disponibles sur la période étudiée, à savoir les particules de diamètre aérodynamique inférieur à 10 µm (PM₁₀), le NO₂, l'O₃ et le SO₂.

b) Exposition domestique

La qualité de l'air intérieur est également un facteur majeur d'exposition en population générale. L'évaluation de son impact sur la santé est plus difficile en raison des sources variées d'exposition.

Impact dans le Monde :

D'après l'OMS (2021), la pollution de l'air domestique serait responsable de 2,9 millions de décès prématurés par an. La pollution de l'air domestique est un problème majeur au niveau mondial, essentiellement à cause du chauffage par biomasse (28).

Impact en France :

En France le chauffage par biomasse est moins fréquent. L'impact d'autres types de pollution domestique comme les composés organiques volatils, les moisissures ou encore les allergènes d'acariens a été estimé (29).

Une première estimation de l'agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) estimait un coût socioéconomique des polluants de l'air intérieur à 19 milliards d'euros annuellement en France (29).

L'évaluation de l'exposition à la pollution de l'air intérieur a été réalisée par un questionnaire dans l'enquête ELISABET. L'exploitation de ce questionnaire pour estimer la qualité de l'air est complexe. Actuellement, un travail est en cours pour estimer l'association entre les questions du questionnaire et l'exposition mesurée par micro-capteur de particules sur un groupe volontaire. Ce travail fait partie d'une thèse en cours et en collaboration avec les Laboratoires d'Optique Atmosphérique et de PhysicoChimie des Processus de Combustion et de l'Atmosphère de l'Université de Lille.

Actuellement la partie du questionnaire la plus pertinente est celle concernant les moisissures du logement de chaque participant (hors cuisine et salle de bain). Une étude antérieure à cette thèse a été réalisée au sein de la population ELISABET et a retrouvé une association entre la présence de moisissures dans l'habitation et une diminution prédite du volume expiratoire maximal en une seconde (VEMS) ainsi qu'une augmentation des sifflements respiratoires (30). Par ailleurs il s'agit d'un marqueur de la qualité des logements connu en lien avec des risques sanitaires (31).

c) Exposition professionnelle

Impact dans le Monde :

D'après l'OMS (2021), environ 2 millions de décès sont liés aux risques professionnels chaque année (32).

Impact en France :

D'après Santé Publique France (données 2006 mises à jour en 2019), environ 20 000 décès seraient attribuables chaque année à des facteurs professionnels (33).

La communauté urbaine de Dunkerque est marquée par une forte industrialisation, ce qui se traduit par une prévalence plus élevée des expositions professionnelles. Par exemple, l'exposition professionnelle à la silice cristalline était 2 fois plus élevée à Dunkerque qu'à Lille (34).

Pour ce travail, qui porte sur la population générale, nous nous sommes intéressés à un marqueur d'exposition professionnelle non spécifique. L'exposition aux vapeurs, poussières, gaz et fumées est largement utilisée comme indicateur d'exposition professionnelle dans les études épidémiologiques portant sur les maladies respiratoires chroniques (35). C'est l'indicateur que nous avons retenu pour ce travail.

L'estimation des expositions professionnelles aux vapeurs, poussières, gaz et fumées dans la population ELISABET a été déterminée à partir d'une matrice emplois-expositions par les médecins du travail du Laboratoire IMPact de l'Environnement Chimique sur la Santé (IMPECS) de l'université de Lille (34) dans le cadre d'une collaboration.

d) Exposition aux métaux lourds

Impact dans le Monde :

D'après l'OMS, l'exposition au plomb entraînerait 1,5 millions de décès (en 2021) (36). L'OMS recommande le seuil de plombémie à 50 µg/L comme seuil déclencheur d'un examen approfondi des sources d'exposition d'une personne au plomb, ainsi que de la mise en place d'actions visant à réduire ou à supprimer cette exposition (37,38).

Impact en France :

En France les niveaux d'expositions restent modestes. Par exemple, dans l'étude Esteban réalisée en population générale entre 2014 et 2026 chez des adultes âgés de 18 à 74 ans, 31,19 % avaient une plombémie supérieure à 25 µg/L, 4,98 % une plombémie supérieure à 50 µg/L (seuil de l'OMS) et 0,62 % une plombémie supérieure à 100 µg/L (39).

Cependant, nous ne pouvons pas exclure les effets des faibles doses.

Dans l'enquête ELISABET, nous avons mesuré divers métaux lourds (plomb, cadmium, chrome, béryllium, cobalt, antimoine, thallium, vanadium, manganèse, nickel, zinc, arsenic, aluminium et mercure).

L'effet de lot (ou effet batch) a fait l'objet d'une mesure simultanée et pouvait avoir un impact statistiquement important sur la mesure de ces métaux. Il s'agit d'une variabilité technique qui nécessite un ajustement statistique (40).

Nous avons retenu deux métaux : le cadmium et le plomb car ce sont des expositions environnementales fréquentes et des liens avec la santé sont connus pour ces métaux (41–43). De plus, la qualité technique de la mesure était acceptable, avec un effet de lot qui pouvait être corrigé par ajustement.

4. Etude des liens entre expositions environnementales et pathologies : rôle des intermédiaires

La figure ci-dessous représente les liens étudiés entre les expositions environnementales disponibles dans ELISABET et différentes variables. Ces études sont détaillées dans les paragraphes suivants.

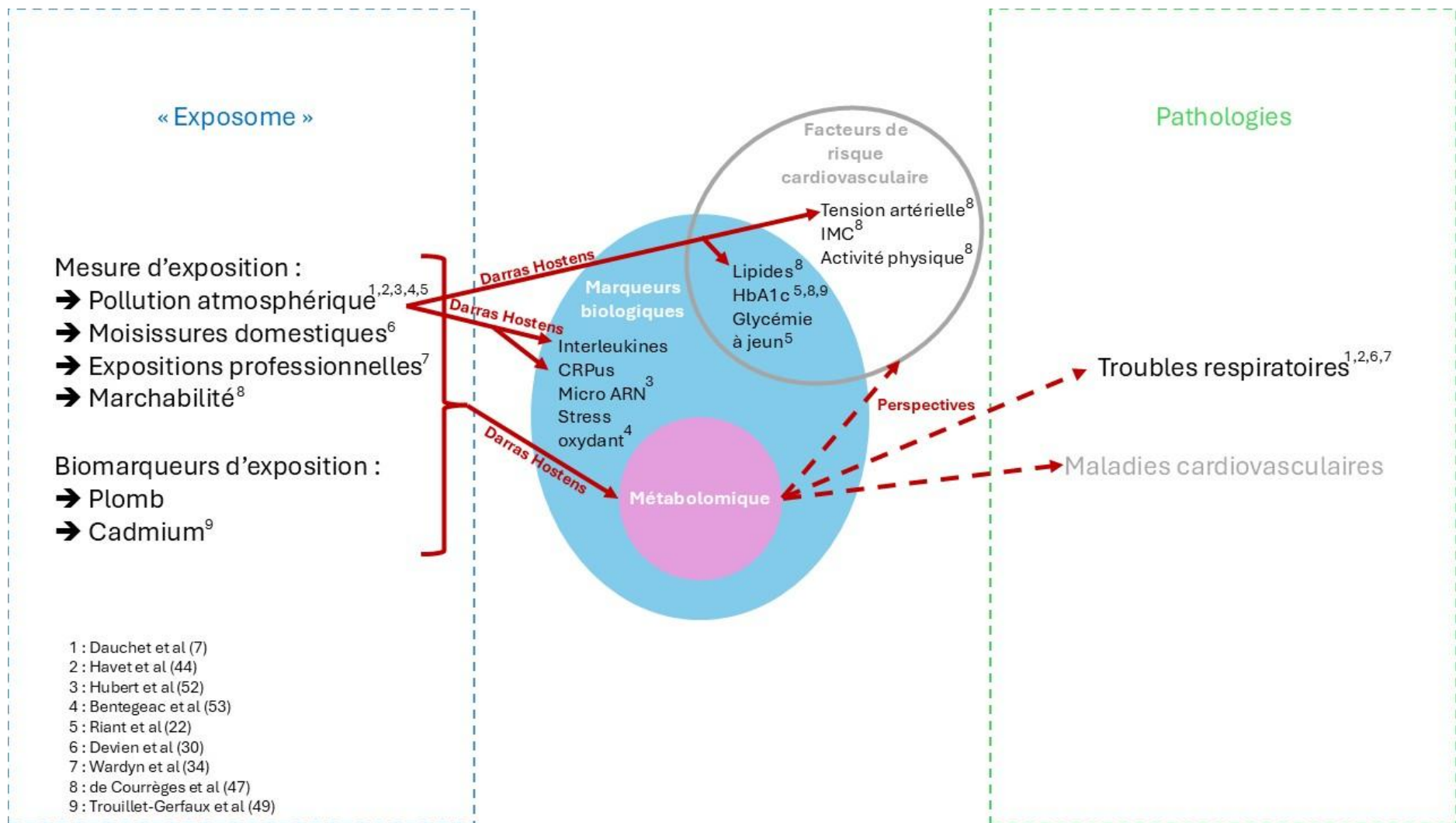


Figure I.A.4.1 : Etudes des liens entre expositions environnementales et différentes variables dans l'enquête ELISABET

a) Etude directe du lien entre exposition environnementale et santé

La quantification des associations entre environnement et santé se fait en étudiant les associations entre ces expositions et des événements de santé. Pour ce qui concerne l'étude ELISABET, des travaux antérieurs à cette thèse (figure I.A.4.1), ont étudié les associations entre la fonction respiratoire et : l'exposition aux moisissures du domicile (30), l'exposition à la pollution atmosphérique à court terme (7) et à long terme (44) et l'exposition professionnelle en population générale aux poussières de silice cristalline (34).

L'avantage de cette méthode est qu'elle permet de quantifier directement la relation effet-dose.

Les inconvénients sont la nécessité de populations importantes, le plus souvent de suivis longs lorsque l'on s'intéresse à des événements incidents comme les cancers et les maladies cardiovasculaires. Ce type d'étude peut ne pas mettre en évidence de petits effets ou des effets à très long terme. Enfin, la question de causalité est difficile à établir puisque cette méthode ne permet pas d'étudier les mécanismes physiopathologiques.

b) Critères intermédiaires ciblés

Pour étudier la relation entre une exposition et une pathologie, il est également possible de passer par l'étude de critères intermédiaires. Ces marqueurs intermédiaires, reliés à la fois à l'exposition et à la pathologie permettent de renforcer les hypothèses de liens entre exposition et pathologie. Certains ont appelé cette approche « meet in the middle » (45).

(1) Facteurs de risque connus

Les variables intermédiaires peuvent être des facteurs de risque connus d'une pathologie. L'exemple typique concerne les facteurs de risque des maladies cardiovasculaires. En effet, les facteurs de risque cardiovasculaire sont mesurés en routine (bilan sanguin, prise de la tension artérielle, calcul de l'IMC, estimation de l'activité physique) et sont des facteurs intermédiaires entre l'exposition environnementale et les maladies cardiovasculaires. Les liens causaux entre ces facteurs de risque et maladies cardiovasculaires, qui font l'objet d'un dépistage et d'une prise en charge en routine, sont bien connus. Ces facteurs de risque sont donc des marqueurs d'effet des expositions environnementales très robustes.

Ce modèle n'est aujourd'hui essentiellement applicable qu'aux maladies cardiovasculaires. L'équivalent de ces facteurs de risque cardiovasculaire n'existant pas pour d'autres pathologies fréquentes comme les cancers (en dehors de l'obésité qui est facteur de risque pour de nombreuses pathologies (46)).

Dans l'enquête ELISABET, un certain nombre de travaux antérieurs à cette thèse (figure I.A.4.1) ont été réalisés sur les liens entre exposition environnementale et ces facteurs de risque cardiovasculaire.

Une étude a regardé les associations entre la marchabilité et les facteurs de risque cardiovasculaire et a retrouvé une augmentation significative de l'activité physique avec la marchabilité et une diminution de la pression artérielle ainsi que de l'IMC (47).

Par ailleurs, deux études ont étudié des relations avec le diabète. Ainsi, une étude a regardé les liens entre l'exposition à long terme à la pollution atmosphérique et l'hémoglobine glyquée (HbA1c), la glycémie à jeun et le diabète. Des associations positives ont été retrouvées entre l'exposition aux PM₁₀ et au NO₂ et l'HbA1c (48). La deuxième étude portait sur les associations entre le cadmium sanguin et l'HbA1c et a montré une augmentation significative de l'HbA1c avec la concentration sanguine en cadmium chez les non-fumeurs et les fumeurs (49).

Enfin, l'étude des facteurs de risque cardiovasculaire fait l'objet d'un des travaux de cette thèse. Celui-ci porte sur les associations entre l'exposition résidentielle à la pollution atmosphérique et les lipides ainsi que la pression artérielle. Ce travail permet de couvrir les investigations sur les facteurs de risque cardiovasculaire dans ELISABET et a été publié en première auteure (50).

(2) Biomarqueurs

La variable intermédiaire peut aussi être un biomarqueur. Un marqueur biologique ou biomarqueur est une caractéristique biologique mesurable reflétant un processus (pathologique ou non) (51).

L'étude ELISABET a mesuré différents biomarqueurs dont certains ont été étudiés dans d'autres travaux distincts de cette thèse (figure I.A.4.1).

Une étude ayant pour objectif d'examiner l'association entre l'expression des micro-ARN et l'exposition résidentielle à long terme à la pollution atmosphérique suggérait que les micro-ARN circulants pourraient être des biomarqueurs des effets de la pollution atmosphérique (52).

Une autre analyse sur l'enquête ELISABET a regardé les associations entre l'exposition à court terme et à long terme à la pollution atmosphérique et des biomarqueurs du stress oxydant. Les résultats n'ont pas montré d'association significative (53).

Enfin, l'étude de biomarqueurs fait l'objet d'un des travaux de cette thèse. Celui-ci porte sur les associations entre l'exposition à court terme et résidentielle à la pollution atmosphérique et des biomarqueurs de l'inflammation de bas grade (CRPus et cytokines). Ce travail permet de couvrir les investigations sur les biomarqueurs mesurés dans ELISABET a été publié en première auteure (54).

L'avantage de cette approche par biomarqueurs est qu'elle permet d'explorer les voies métaboliques connues ou pour lesquelles on a des hypothèses fortes.

(3) Avantages et inconvénients des intermédiaires ciblés

L'approche par facteurs de risque connus ou par biomarqueurs constitue dans les 2 cas une approche ciblée.

Cette approche nécessite de formuler une hypothèse a priori sur cette relation.

L'approche ciblée présente plusieurs avantages : l'hypothèse posée est claire, l'interprétation des résultats est assez simple et le jeu de donnée est de petite taille (par rapport aux approches

non ciblées). De plus, le nombre d'hypothèses étant restreint, cette approche ne nécessite pas ou nécessite peu de correction pour les tests multiples, ce qui permet de conserver la puissance statistique.

Cependant, l'approche ciblée présente plusieurs inconvénients. En effet, elle est basée sur des hypothèses biomédicales et donc contrainte par la connaissance que l'on peut avoir a priori des relations entre chaque exposition et la physiopathologie. Elle se limite ainsi à un nombre restreint de combinaisons possibles et de paramètres cliniques (phénotypes). Par ailleurs, si l'hypothèse de départ est fautive, alors les résultats seront ininterprétables.

c) Critères intermédiaires non ciblés

L'approche dite « omique » permet de mesurer des variables intermédiaires de manière non ciblée. Ces variables intermédiaires sont appelées « signaux ».

Cette approche comprend par exemple, la génomique qui vise à étudier l'ensemble des génomes, la transcriptomique qui s'intéresse à l'ensemble des ARN, la protéomique qui étudie l'ensemble des protéines et la métabolomique qui étudie l'ensemble des métabolites (55).

Dans ce travail, nous nous intéressons aux métabolites comme intermédiaires dans les relations entre les expositions environnementales et les phénotypes. Nous nous intéressons donc à la métabolomique non-ciblée.

B. Apports de la métabolomique

1. Définition de la métabolomique

La métabolomique se définit comme l'analyse exhaustive et semi-quantitative de l'ensemble des métabolites mesurables dans un échantillon biologique (56,57).

La métabolomique est le signal cellulaire ultime, en bout de chaîne de l'activité cellulaire, derrière la protéomique et la génomique (58). Elle permet de détecter précocement des perturbations biologiques (58).

De plus le profil métabolomique représente l'aboutissement de la réponse cellulaire et constitue un lien central entre les gènes et les phénotypes (59,60).

2. Méthodes de mesures

Deux principales techniques analytiques sont disponibles pour la métabolomique : la spectroscopie de résonance magnétique nucléaire (RMN) ou la spectrométrie de masse (SM) (61). La RMN est une technique qui permet de mesurer un nombre limité de métabolites — généralement entre 40 et 200 (62). Couplée à une technique de chromatographie, la SM présente plusieurs avantages par rapport à la RMN : elle est plus sensible et peut détecter un plus grand nombre de métabolites, y compris ceux présents en faibles quantités (61). Ainsi, la chromatographie couplée à la spectrométrie de masse (CSM) permet la mesure de plusieurs centaines, voire de plusieurs milliers de métabolites, et tend à devenir la technique la plus utilisée (63). Étant donné que la CSM permet de mesurer un plus grand nombre de

métabolites, l'exploration des voies métaboliques peut être plus détaillée, le nombre de métabolites quantifiés par voie étant plus élevé.

Cependant, la CSM présente aussi certains inconvénients par rapport à la RMN, notamment sa complexité et la difficulté de son application à grande échelle. Si des études portant sur quelques centaines à quelques milliers d'individus peuvent être réalisées avec la CSM (64–66), on reste encore loin des méga-cohortes fondées sur la RMN, qui incluent plusieurs dizaines de milliers de participants (61). En outre, l'analyse statistique des données métabolomiques issues de la CSM en épidémiologie environnementale demeure particulièrement complexe, en raison du grand nombre de variables.

La chromatographie liquide ou gazeuse à ultra-haute performance couplée à la spectrométrie de masse a été développée au début des années 2000 (67).

Cette technique permet d'obtenir, pour chaque signal, une caractéristique fondée sur le temps de rétention (obtenu par chromatographie) et le rapport masse/charge (obtenu par spectrométrie de masse). Ces deux paramètres sont ensuite comparés à des standards ou à des bases de données de référence afin de tenter d'identifier (annoter) chaque métabolite (68). Parmi les bases de données disponibles, on peut citer la *Human Metabolome DataBase* (HMDB) (69,70), ou encore *Metabolite Link* (Metlin) (71).

Les métabolites sont annotés avec des degrés variables de certitude. Ainsi, un métabolite annoté au niveau 1 correspond à un métabolite formellement identifié. Au niveau 2, il s'agit d'un métabolite annoté de façon putative. Au niveau 3, on parle d'une classe de composés caractérisée de manière putative, tandis qu'un métabolite est considéré comme inconnu lorsque le niveau d'annotation est de 4 (72). A partir de l'annotation des métabolites, un identifiant peut leur être attribué. La base de données Chemical Entities of Biological Interest (ChEBI) répertorie les métabolites en selon leur identifiant attribué (73).

Par ailleurs, les métabolites appartiennent à des voies métaboliques. Il existe donc des bases de données permettant de classer les métabolites dans une ou plusieurs voies auxquelles ils appartiennent (74), par exemple la *Kyoto Encyclopedia of Genes and Genomes* (KEGG) (75,76). Lorsqu'on explore les associations entre une exposition environnementale et les métabolites, il est alors possible d'étudier quelles voies métaboliques sont impliquées dans ces associations. Une voie métabolique est définie comme une série de réactions chimiques conduisant à la transformation d'une molécule en un produit (métabolite) (77,78).

L'approche la plus fréquente pour retrouver les voies métaboliques consiste à rechercher un enrichissement en voies métaboliques, notamment à l'aide d'une méthode appelée *over-representation analysis* (ORA). Celle-ci repose sur des outils bioinformatiques tels que *MetaboAnalyst*, qui permet d'entrer une liste de métabolites d'intérêt (c'est-à-dire significativement associés à l'exposition), puis l'algorithme cartographie les voies métaboliques impliquées et génère une analyse de sur-représentation des voies ainsi qu'une analyse de topologie des voies (79).

Il convient de noter que les associations entre une exposition et des métabolites non annotés peuvent être difficiles à interpréter (80). Cela pose la question de la réalisation d'analyses statistiques uniquement sur les métabolites annotés afin de limiter le nombre de tests statistiques et donc la correction pour tests multiples et de conserver ainsi la puissance statistique.

3. Place de la bio-informatique

Si les études métabolomiques les plus anciennes se limitaient à comparer les résultats biologiques des sujets exposés et non exposés dans une perspective de classification, les développements récents de la bio-informatique ont permis d'enrichir considérablement l'interprétation des signatures biologiques, grâce à l'exploitation automatisée des connaissances des voies métaboliques et des réseaux moléculaires disponibles dans les bases de données biomédicales publiques et les publications (81). L'analyse bio-informatique des réseaux métaboliques est un domaine en pleine expansion de la biologie des systèmes. Elle repose sur le fait que les métabolites associés à une exposition ne sont pas indépendants, car contraints par un système complexe d'interactions moléculaires constitué par l'ensemble des réactions biochimiques pouvant avoir lieu au sein d'une cellule ou d'un tissu. Ces milliers de réactions (plus de 13 000 pour la dernière base de données du métabolisme humain (82)) permettent de cartographier, et donc de contextualiser, les modifications métaboliques liées à une exposition. Les algorithmes sur les réseaux, dont certains inspirés des méthodes utilisées dans les réseaux sociaux, permettent d'identifier au sein de ces structures complexes les voies métaboliques spécifiquement modulées (83). Ainsi, l'analyse des caractéristiques métaboliques produites par les mesures chimiques non ciblées devrait permettre d'identifier les voies moléculaires (et non plus les seules molécules) éventuellement dysrégulées ou altérées par une exposition environnementale.

4. Les avantages de la métabolomiques

La métabolomique comporte plusieurs avantages.

Tout d'abord c'est une approche non ciblée, qui ne nécessite pas d'hypothèses préalables et permet de rechercher des liens entre expositions environnementales, sans hypothèse a priori. Les expositions à des doses faibles ou modérées peuvent avoir des effets subcliniques entraînant des impacts à long terme sur la santé. Ces impacts à long terme sont possiblement difficiles à étudier car ils peuvent évoluer silencieusement durant de nombreuses années.

Il est important d'étudier les liens entre les expositions environnementales et les modifications phénotypiques afin de mieux comprendre les effets et leurs mécanismes biologiques et de pouvoir ainsi proposer des recommandations de santé publique.

Elle permet aussi d'utiliser l'ensemble des données, limitant ainsi le risque de biais de publication.

Un autre avantage est qu'elle offre des possibilités de détecter de nouveaux signaux non anticipés et d'obtenir ainsi des connaissances innovantes (58).

Cette approche est potentiellement très sensible et permet de détecter des modifications métaboliques très précoces (par exemple pour explorer la toxicité dans les études expérimentales) (60).

5. Les inconvénients de la métabolomique

L'un des inconvénients de la métabolomique est la difficulté d'interprétation des résultats. Les modifications métaboliques peuvent être des adaptations normales (physiologiques), sans conséquence pathologique, voire plutôt des effets bénéfiques. Il est difficile de

distinguer des perturbations traduisant un impact sanitaire des perturbations relevant d'adaptations physiologiques du métabolisme (58).

Par ailleurs, il existe une variabilité interindividuelle et intra individuelle du métabolome en raison de multiples facteurs confondants (mode de vie, alimentation, état pathologique, etc.) (58,60).

Cette approche est apparentée à la « fishing expedition » puisqu'elle mesure tous les métabolites mesurables dans l'échantillon biologique, rendant complexe le sens à donner aux résultats a posteriori (58).

Par ailleurs, un autre inconvénient provient de la taille importante des jeux de données produits en métabolomique conférant une haute dimensionnalité (84,85).

Cette taille importante des jeux de données induit une problématique liée aux tests multiples.

Le principal problème lié aux tests multiples est celui de l'inflation du taux d'erreur de type I (faux positifs). Lorsque de nombreuses comparaisons sont effectuées, la probabilité d'obtenir au moins un faux positif (résultat significatif alors qu'il n'y a pas d'association) augmente. Cela signifie que même si toutes les comparaisons individuelles sont correctes, il y a une forte probabilité d'obtenir des résultats statistiquement significatifs par pur hasard. Les corrections pour tests multiples sont utiles et conservatrices, mais elles peuvent parfois se révéler trop strictes, entraînant une perte de puissance statistique et un risque accru de faux négatifs.

6. Les usages de la métabolomique

La métabolomique a différentes applications.

Dans cette thèse, on s'intéresse à l'étude des modifications biologiques liées aux expositions environnementales courantes en population générale. On cherche à explorer la place de la métabolomique dans ce contexte particulier.

La métabolomique a d'autres applications en santé environnement qui n'entrent pas directement dans le champ de cette thèse.

Elle permet l'évaluation de la toxicité de substances en conditions expérimentales notamment en expérimentation animale. Elle peut ainsi découvrir des biomarqueurs candidats d'exposition. C'est biomarqueurs ou cette combinaison de biomarqueurs permettraient d'évaluer une exposition environnementale par un simple échantillon biologique (60).

La métabolomique peut être utilisée pour rechercher des biomarqueurs précoces de pathologies, par exemple les maladies cardiovasculaires ou les cancers (58,60). Ce point n'entre pas dans ce travail de thèse mais est prévu dans la suite des analyses de METALISABET dans le cadre d'étude de médiation avec les TVO et les facteurs de risque cardiovasculaire.

Un autre enjeu majeur est l'identification de proxies (c'est-à-dire des indicateurs indirects) d'exposition ou de maladies, susceptibles de constituer des outils de prévention, de diagnostic ou de suivi (86).

Enfin, la métabolomique est fréquemment utilisée dans un but de compréhension mécanistique, par l'étude des métabolites associés à un phénotype biologique afin d'apporter un éclairage sur les mécanismes biologiques sous-jacents (87). C'est le cadre de ce travail de thèse. La métabolomique offre ainsi la possibilité d'explorer les voies métaboliques impactées par l'exposition environnementale étudiée. La recherche de voies métaboliques est essentielle pour pouvoir formuler des hypothèses mécanistiques (60).

Un des intérêts est de mettre en évidence des perturbations métaboliques infracliniques liées à une exposition, confirmant ainsi ou suggérant de nouvelles hypothèses sur les mécanismes physiopathologiques sous-jacents (69).

7. Enjeux méthodologiques de l'analyse de données de métabolomique

L'étude statistique des associations entre exposition environnementale et métabolomique en population générale présente des difficultés méthodologiques spécifiques.

Lorsque l'on travaille sur des données de grande taille, il est nécessaire de réfléchir à une adaptation des méthodes statistiques.

De nombreuses études métabolomiques ont été menées sur de petits échantillons de population (quelques dizaines de sujets). Par exemple, une étude a examiné les associations entre la pollution atmosphérique et le métabolisme des acides aminés chez 73 adultes en Chine à l'aide d'une analyse discriminante par moindres carrés partiels orthogonale (*orthogonal partial least squares discriminant analysis*, O-PLS-DA) (88). Une autre étude, réalisée chez 92 jeunes étudiants en Chine, a utilisé l'analyse en composantes principales (ACP) et l'O-PLS-DA pour explorer les associations entre les particules en suspension (PM), les bioaérosols et la métabolomique (89). Enfin, une étude interventionnelle menée auprès de 10 étudiants en bonne santé au Brésil a également eu recours à l'ACP et à une analyse discriminante par moindres carrés partiels (*partial least squares discriminant analysis*, PLS-DA) afin d'examiner les associations entre la pollution liée au trafic et le profil métabolomique sérique (90).

Dans ce type de configuration, où le nombre de variables dépasse largement celui des sujets (haute dimension), seules des analyses multidimensionnelles sont envisageables. Les méthodes multidimensionnelles sont des méthodes utilisées pour étudier simultanément de nombreuses variables en réduisant le nombre de dimensions, dans un but de résumer les informations. Les ACP permettent de réduire le nombre de variables en un nombre limité de composantes ou dimensions. L'inconvénient est que ces dimensions sont définies indépendamment de l'exposition. Les profils ainsi définis peuvent ne pas être corrélés aux métabolites les plus associés à l'exposition, ce qui peut masquer des associations. D'autres méthodes comme la régression des moindres carrés partiels (*partial least square*, PLS), prennent en compte l'exposition dans la définition des dimensions (60), ce qui permet d'identifier des associations même avec des effectifs réduits. Ce type d'analyse ne permet toutefois pas de tester formellement les hypothèses H0 et H1 avec un test statistique.

Les progrès techniques récents et la baisse des coûts permettent désormais de réaliser des études sur des échantillons de plus grande taille. L'augmentation de la taille d'échantillon élargit les possibilités d'analyses statistiques : les analyses par métabolite individuel, avec corrections pour tests multiples, deviennent alors envisageables.

Par ailleurs, une base de données métabolomique de grande taille permet une variabilité importante dans les méthodes statistiques employées, ce qui complexifie les comparaisons entre études et limite l'évaluation de la reproductibilité des résultats. Enfin, l'augmentation des effectifs pose également la question de la place des analyses multidimensionnelles.

Pour ce travail, nous disposons de données métabolomiques sur un échantillon d'un millier d'individus. Notre enjeu était donc d'analyser un jeu de données de grande taille, et il était nécessaire de réfléchir à la stratégie d'analyse. En effet, nous nous attendons, par exemple, à des effets faibles ou modérés de l'exposition à la pollution sur le métabolisme, il est donc nécessaire d'utiliser des statistiques permettant de préserver la sensibilité de l'analyse tout en évitant le risque de faux positifs lié au tests multiples.

À l'heure actuelle, il n'existe pas de recommandations établies concernant les analyses statistiques à utiliser pour étudier les associations entre expositions environnementales et données de métabolomique issues de la chromatographie-spectrométrie de masse.

Il est toutefois nécessaire d'harmoniser les méthodologies statistiques et la présentation des résultats afin de permettre la comparabilité entre les études et de prévenir le phénomène de *p-hacking* qui consiste à adapter le choix des méthodes statistiques en fonction des résultats obtenus, voire multiplier les analyses jusqu'à l'obtention de résultats significatifs.

C. Objectif

La question scientifique de cette thèse est donc : quelle démarche utiliser pour améliorer la compréhension mécanistique des associations observées entre des expositions environnementales courantes en population générale et l'impact sur la santé.

Le plan de travail est donc le suivant :

- 1) Etudier les relations entre exposition et pathologie à l'aide d'un intermédiaire concernant les facteurs de risque cardiovasculaire.
- 2) Etudier ces associations avec un intermédiaire concernant des biomarqueurs.
- 3) Faire une revue de la littérature afin de connaître la place de la métabolomique dans ce contexte de population générale. C'est-à-dire, recenser les méthodologies statistiques utilisées dans la littérature (portant sur les liens entre environnement et métabolomique) afin d'en déduire une stratégie d'analyse.
- 4) Etudier les associations entre l'ensemble des expositions connues dans ELISABET et les modifications métaboliques.
- 5) Etudier de manière approfondie une exposition (au plomb) où les associations sont fortes afin d'exploiter la réflexion sur la démarche d'interprétation des voies métaboliques.

II. METHODE GENERALE

A. Matériel

1. Présentation de l'enquête ELISABET

a) Recrutement des participants

L'Enquête Littoral Souffle Air Biologie Environnement (ELISABET) est une étude transversale réalisée dans le nord de la France (Lille et Dunkerque). Elle a inclus 3175 hommes et femmes âgés de 40 à 65 ans, résidant dans la même ville ou sa communauté urbaine depuis au moins 5 ans (soit Lille, soit Dunkerque). L'objectif principal de l'enquête ELISABET était de comparer la prévalence des TVO entre la communauté urbaine de Dunkerque, fortement exposée à la pollution industrielle, et la communauté urbaine de Lille, moins industrialisée.

Les critères de non-inclusion étaient les suivants : les femmes enceintes, les personnes privées de liberté ou sous protection légale, les personnes hors d'état d'exprimer leur consentement et les personnes ne bénéficiant pas d'assurance sociale.

Le recrutement s'est déroulé entre janvier 2011 et novembre 2013.

Les participants ont été sélectionnés à partir des listes électorales, par tirage au sort, avec une stratification selon le sexe, l'âge et la communauté urbaine (de Lille ou de Dunkerque). Ils ont été contactés dans un ordre aléatoire. Pour les communes de petite taille de l'agglomération lilloise, un échantillonnage en grappes a été réalisé pour que la distribution des données démographiques soit la plus homogène possible au sein de l'étude.

Chaque participant sélectionné a reçu une lettre l'invitant à contacter l'équipe de coordination pour convenir d'un rendez-vous de collecte de données. En cas d'absence de réponses, les investigateurs envoyaient des relances par courrier et, lorsque cela était possible, par téléphone.

Le recueil des données a eu lieu au domicile des participants, ou plus rarement, lors d'une consultation dans un centre de santé. Dans tous les cas, un(e) infirmier(ère) diplômé(e) d'État formé(e) a administré un questionnaire détaillé (7,91).

Des examens ont été réalisés, au cours de la même visite : des prélèvements biologiques (sanguins et urinaires), et des prélèvements de cheveux. Un recueil de condensats dans l'air expiré a également été réalisé. Différentes mesures ont été effectuées, comme la mesure de la fréquence cardiaque, de la tension artérielle, la mesure du monoxyde d'azote (NO) dans l'air expiré et dans l'air ambiant, et la mesure du CO dans l'air expiré. Enfin, des tests de la fonction respiratoire (une spirométrie et la mesure de la fraction expirée du monoxyde d'azote (FeNO)) ont été faits.

Le protocole de l'étude (identifiant ClinicalTrials.gov : NCT02490553) a été approuvé par le comité de protection des personnes Nord-Ouest IV (Lille, France ; référence : 2010-A00065-34), conformément à la législation française sur la recherche biomédicale. Tous les participants ont donné leur consentement éclairé écrit pour participer à l'étude.

b) Mesures des expositions

Nous disposons de mesures d'expositions différentes. En effet, nous avons des mesures concernant la pollution atmosphérique : PM₁₀, NO₂, O₃ et SO₂. De plus, nous disposons de mesures de l'exposition domestique aux moisissures. Par ailleurs, des mesures d'expositions professionnelles ont été obtenues : gaz, vapeurs, fumées, poussières et silice cristalline. Enfin, nous avons mesuré les concentrations sanguines en métaux lourds : le cadmium et le plomb.

Ces mesures ont été obtenues pour Lille et Dunkerque (sauf pour le SO₂ uniquement à Dunkerque) ; ces deux zones présentent un trafic automobile important, et la partie nord-ouest de l'agglomération dunkerquoise comporte une zone industrielle.

c) Pollution atmosphérique

Nous avons mesuré les polluants atmosphériques suivants : PM₁₀, NO₂, O₃ et SO₂. Pour les PM₁₀ et le NO₂ nous avons l'exposition à court terme et long terme. Pour l'O₃ nous avons l'exposition à court terme uniquement et pour le SO₂, nous avons l'exposition à long terme uniquement (et uniquement pour la communauté urbaine de Dunkerque). L'exposition à court terme a été définie comme étant la moyenne journalière (du jour et de la veille des prélèvements sanguins et urinaires) d'après les mesures des stations de chaque communauté urbaine (de Lille ou de Dunkerque). L'exposition à long terme a été définie comme étant la moyenne annuelle selon l'adresse résidentielle du participant et son année d'inclusion dans l'étude (modélisation Atmospheric Dispersion Modelling System (ADMS)). Les données d'exposition ont été fournies par Atmo Haut de France (92,93). Nous parlerons d'exposition résidentielle lorsque nous parlerons de l'exposition à la pollution atmosphérique qui rend compte d'une exposition à long terme.

Entre 2010 et 2013 à Lille et entre 2012 et 2013 à Dunkerque, les concentrations annuelles moyennes de NO₂, de PM₁₀ et (pour Dunkerque uniquement) de SO₂ ont été estimées par l'organisme de surveillance de la qualité de l'air ATMO Hauts-de-France (Lille, France) (48) à l'aide d'un système de modélisation de la dispersion atmosphérique (94).

Nous avons utilisé une résolution spatiale de 25 × 25 m et évalué l'exposition annuelle de chaque résidence à l'intérieur de cette grille. Le niveau d'exposition global à la pollution pour chaque participant a été calculé comme la moyenne des concentrations de polluants dans l'année de son inclusion dans l'étude et dans les années précédentes.

d) Exposition domestique

Concernant l'exposition domestique, nous avons interrogé les participants sur la présence de moisissures visibles dans les différentes pièces de l'habitation (30). Nous avons gardé la présence de moisissures (oui ou non) dans le logement hors cuisine et salle de bains.

e) Exposition professionnelle

Nous avons évalué les expositions professionnelles aux gaz, vapeur, fumées et poussières en interrogeant les participants sur leur cursus laboris (trajectoire professionnelle). Les participants ont été interrogés sur leur premier emploi, leur dernier emploi et l'emploi exercé

le plus longtemps au cours de la carrière, afin d'identifier les emplois susceptibles d'avoir entraîné une exposition aux vapeurs, poussières, gaz et fumées. La collaboration avec les médecins du travail de Lille a permis la création d'une matrice emploi-exposition puis d'un indice d'exposition cumulée aux « vapeurs-poussières-gaz-fumées » (34).

Par ailleurs, l'exposition de chaque participant à la silice cristalline a été quantifiée à l'aide d'un indice cumulatif d'exposition (CEIsilica), calculé pour chaque emploi occupé et pour l'ensemble de la carrière, à partir de la matrice emploi-exposition (JEM) française Matgène élaborée par Santé Publique France (34).

f) Expositions aux métaux lourds

L'exposition au cadmium et au plomb a été évaluée en mesurant leur concentration sanguine pour chaque participant. Les niveaux de cadmium et de plomb ont été déterminés à partir de sang veineux prélevé dans des tubes spécifiques aux oligo-éléments (Vacutainer Trace Elements K2EDTA 10,8 mg, réf. 368381, bouchon bleu, Becton Dickinson, Le Pont de Claix, France), puis conservés rapidement à 4 °C. Les échantillons de sang total ont ensuite été dilués au ratio 1:10 dans une solution aqueuse contenant 0,1 % de Triton et 0,05 % d'ammoniaque. Les échantillons préparés ont été analysés sans délai à l'aide d'un spectromètre de masse à plasma à couplage inductif doté d'une interface de collision/réaction (Varian 820-MS, Bruker, Wissembourg, France) (49).

2. Présentation de Metalisabet

Le projet Metalisabet financé par le contrat numéro 21ESD0019 entre l'Institut Pasteur de Lille, le CEA (Commissariat à l'Énergie Atomique et aux Énergies Alternatives), l'INRAE (Institut National de la Recherche pour l'Agriculture, l'Alimentation et l'Environnement) et l'ADEME (Agence de l'Environnement et de la Maîtrise de l'Énergie), se concentre sur l'impact de la pollution atmosphérique et environnementale sur la santé. Il utilise une approche multidisciplinaire combinant épidémiologie, chimie analytique et bio-informatique pour explorer les relations entre les expositions environnementales et les perturbations métaboliques et endocriniennes.

Les équipes associées au projet Metalisabet sont l'Institut Pasteur de Lille (UMR RID AGE – facteur de risque et déterminants moléculaires des maladies liées au vieillissement), le CEA (UMR CEA-INRAE Université Paris-Saclay Médicaments et Technologies pour la Santé (MTS)) et l'INRAE (UMR ToxAlim 1331 Toulouse).

Le projet Metalisabet vise à analyser les métabolites urinaires de sujets exposés à différents niveaux de pollution atmosphérique ainsi qu'à des environnements domestiques et professionnels variés.

L'étude repose sur les données cliniques, épidémiologiques (expositions) et biologiques de l'enquête ELISABET.

Des échantillons d'urine de l'ensemble des participants à l'enquête (5 ml par sujet, jamais décongelés) étaient disponibles. Ils étaient conservés en plusieurs aliquots dans des

congélateurs indépendants à -80 °C, sous alarme, au Centre de Ressources Biologiques (CRB) de l'Institut Pasteur de Lille (Plateforme CRB nationale, agréée par l'INSERM, dernier agrément au 1/1/2021). Le choix d'échantillons d'urine par rapport au sang a été fait car les échantillons d'urines étaient disponibles en plus grande quantité que les échantillons de sang dans ELISABET. De plus, la concentration des métabolites est amplifiée par le stockage vésiculaire (95).

Pour Metalisabet, les échantillons de 1 000 participants non-fumeurs ont été sélectionnés aléatoirement, stratifiés selon le sexe, la communauté urbaine d'inclusion (communauté urbaine de Lille ou de Dunkerque) et de la date de prélèvement, afin de minimiser les sources de variation. Ces échantillons ont été préparés et transférés au CEA pour l'analyse métabolomique. Seuls les non-fumeurs ont été inclus afin d'éviter une interaction entre le tabagisme (en cours ou ancien) et les expositions environnementales.

Les métabolites urinaires ont été mesurés par deux méthodes complémentaires de chromatographie liquide ultra-haute performance (phase inverse et HILIC, *Hydrophilic Interaction Liquid Chromatography*), couplée à un spectromètre de masse à haute résolution (Q-Exactive, Thermo Fisher Scientific) opéré en modes d'ionisation électrospray positif et négatif, au CEA (Unité MTS, Saclay). La méthode utilisée correspondant donc à la chromatographie liquide couplée à la spectrométrie de masse en tandem à haute résolution (Liquid Chromatography – High Resolution Tandem Mass Spectrometry (LC-HRMS/MS)).

Le traitement des données incluait la détection automatique des signaux, leur réaligement dans les domaines de masse et de temps de rétention, leur intégration, puis la production d'une matrice de données, à l'aide du logiciel XCMS intégré dans la plateforme Workflow4Metabolomics (W4M). Les signaux ont ensuite été annotés à partir de bases de données publiques et des référentiels internes du laboratoire MTS, selon les protocoles développés par l'équipe (96). Les niveaux d'annotation sont fixés par homologie avec un standard pur analysé dans les mêmes conditions (LC-HRMS/MS) et sont codés par des lettres : « a » correspond à la masse exacte (spectrométrie de masse), « b » correspond au temps de rétention et « c » correspond à un spectre MS/MS cohérent.

Les acquisitions ont été réalisées sur trois lots distincts, en mobilisant deux spectromètres de masse différents et deux méthodes chromatographiques, garantissant ainsi la qualité des résultats. Des contrôles qualité systématiques ont été menés, suivis d'une étape essentielle de normalisation inter-batch. Le pré-traitement des données a impliqué des vérifications rigoureuses et des ajustements réalisés par l'équipe de science des données du laboratoire.

B. Méthode de la thèse : stratégie des analyses

Le plan d'analyse pour le travail de thèse est le suivant :

1. Etudier les relations entre exposition et pathologie à l'aide d'un intermédiaire concernant les facteurs de risque cardiovasculaire

Tout d'abord, nous avons utilisé une approche par intermédiaires connus pour étudier les liens entre exposition et santé. Pour cela, nous avons étudié les associations entre l'exposition résidentielle à la pollution de l'air et les facteurs de risque cardiovasculaire chez des adultes vivant dans le nord de la France.

2. Etudier ces associations avec un intermédiaire concernant des biomarqueurs.

Ensuite, nous avons utilisé une approche par intermédiaires constitués de biomarqueurs. Pour cela, nous avons étudié les associations entre l'exposition à court terme et résidentielle à la pollution de l'air et les niveaux de biomarqueurs inflammatoires chez des adultes vivant dans le nord de la France.

3. Revue méthodologique métabolomique

Puis, nous avons fait une revue systématique de la littérature. Plus précisément, nous avons fait une analyse critique des méthodologies statistiques utilisées pour évaluer les associations entre l'exposition environnementale et les variables métabolomiques dans les études épidémiologiques de grande taille (c'est-à-dire des effectifs d'au moins 500 individus). Cette revue nous a permis de proposer une méthodologie adaptée.

4. Analyses métabolomiques

A partir des résultats de la revue de la littérature, nous avons étudié les liens entre les expositions environnementales disponibles dans notre base de données ELISABET et les métabolites urinaires. La analyses métabolomiques sont précisées dans la figure II.B.4.1

Constitution de l'enquête ELISABET (2011-2013)
IMPECS et RID-AGE Lille
Investigateur principal : Luc Dauchet

Projet Metalisabet : Institut Pasteur de Lille, CEA de Paris, INRAE de Toulouse

Analyse des échantillons d'urine (2022) : chromatographie liquide ultra-haute performance – spectrométrie de masse à haute résolution

10 043 signaux

327 métabolites annotés *

Toute exposition

Plomb

Voies métaboliques associées :
Rapport de vraisemblance
Z score
MSEA (uniquement sur le plomb)

Métabolites associés aux expositions

Voies métaboliques associées :
ORA

Interprétation physiopathologique

Preuve de concept :
Variables de démonstration

Réalisé en amont du travail de thèse par Luc Dauchet, l'équipe du CEA et de l'INRAE

Travail de thèse réalisé par Marion Darras-Hostens

* Annotation basée sur 3 critères (temps de rétention, masse mesurée, spectre MS/MS) par homologie avec un standard pur analysé dans les mêmes conditions

Figure II.B.4.1 : Schéma du workflow sur l'approche métabolomique

5. **Analyses développées sur l'exposition au plomb**

Enfin, nous nous sommes intéressés à l'association entre la plombémie et les métabolites urinaires, que nous avons développée plus en détail, afin de d'engager une réflexion sur la démarche d'interprétation de résultat significatifs.

Une analyse approfondie des autres expositions environnementales est prévue pour la suite des travaux.

III. LOTS DE TRAVAIL (work packages)

A. Etudier les relations entre exposition et pathologie à l'aide d'un intermédiaire concernant les facteurs de risque cardiovasculaire.

L'article publié est disponible en **annexe 1**.

1. Introduction

L'exposition aux particules fines (ou particules en suspension) (PM) dans l'air augmente le risque d'événements cardiovasculaires et de décès d'origine cardiovasculaire (97). La pollution atmosphérique pourrait également influencer les facteurs de risque cardiovasculaires en altérant le métabolisme lipidique et en perturbant la pression artérielle. L'exposition aux PM réduit la capacité antioxydante des lipoprotéines de haute densité (HDL), augmentant ainsi la taille des lésions artérioscléreuses (98,99). Une exposition chronique aux polluants atmosphériques provoque un stress oxydatif, responsable de modifications de l'endothélium, du métabolisme lipidique et de la progression de la plaque artérioscléreuse (100).

Cependant, les données épidémiologiques à ce sujet sont discordantes. Dans une revue systématique, Gaio *et al.*, (101) ont étudié les associations entre pollution atmosphérique et biomarqueurs lipidiques (HDLs, triglycérides (TGs), cholestérol total (CT) et lipoprotéines de basse densité (LDLs) ou dyslipidémie. Ils ont inclus 22 études, dont 8 études de cohorte et 14 études transversales. La moitié des études portaient sur la population générale, tandis que l'autre moitié ciblait des populations spécifiques. La majorité des études provenait des États-Unis et d'Asie (respectivement 9 et 7 études), 2 provenaient du Moyen-Orient et seulement 4 d'Europe de l'Ouest.

Concernant les associations entre l'exposition à long terme (plus de 30 jours) à la pollution et les niveaux lipidiques, 14 études ont trouvé une association avec le CT, 13 avec les HDLs, 9 avec les LDLs et enfin 15 ont trouvé une association avec les TG. Parmi les études avec un échantillon grande taille, Cai *et al.*, ont mis en évidence une association entre l'exposition aux PM₁₀ et au NO₂ et une augmentation des TG chez 114 082 personnes issues de la population générale en Norvège et aux Pays-Bas (102). Une autre étude transversale menée auprès de 100 867 adultes en Corée du Sud a montré une association positive entre l'exposition aux PM₁₀ et au NO₂ et la dyslipidémie (103).

D'autres études incluses dans la revue se sont intéressées à des populations spécifiques. Par exemple, une étude transversale menée auprès de patients (n = 6 587), consultant pour une coronarographie en Caroline du Nord, États-Unis a trouvé des associations entre l'exposition à long terme aux PM et les LDLs, le CT et les TGs (104). Seules 3 de ces études ont pu être incluses dans une méta-analyse, qui a montré une association entre l'exposition aux PM₁₀ et au NO₂ et les TGs. Toutefois, le faible nombre d'études incluses dans la méta-analyse a conduit les auteurs à conclure qu'il fallait davantage de preuves.

Depuis, d'autres études épidémiologiques sur les associations entre la pollution et les lipides ont été publiées. Par exemple, certaines études ont montré une association négative entre l'exposition aux PM et les taux de HDL sanguins dans des populations spécifiques : une étude menée chez 12 778 soldats coréens exposés à long terme aux PM_{2,5} (moyenne journalière = 19,6 µg/m³) (105) et une étude menée chez 88 étudiants universitaires en bonne santé en Chine exposés à court terme aux PM (concentration médiane des PM_{2,5} = 70,22 µg/m³) (106). Une autre étude chinoise réalisée auprès d'enfants et d'adolescents (n = 12 814) a trouvé une association significative entre l'exposition aux PM₁₀ (concentration journalière moyenne = 104

$\mu\text{g}/\text{m}^3$) et au NO_2 (concentration journalière moyenne = $40 \mu\text{g}/\text{m}^3$) et le CT, mais aucune association significative avec les TGs ou le HDL (107). Cette hétérogénéité suggère que le profil lipidique pourrait être impacté par la pollution atmosphérique selon des mécanismes encore mal caractérisés.

Par ailleurs, la relation entre pollution atmosphérique et pression artérielle reste insuffisamment claire, comme le souligne la conclusion de la revue systématique menée par Yang *et al.*, (108). Cette revue a inclus 65 articles regroupant 100 études, et une méta-analyse a été réalisée sur l'ensemble de ces 100 études. Les études provenaient de 16 pays différents, mais moins d'un quart de ces 100 études étaient situés en Europe de l'Ouest. Les études étaient transversales, longitudinales, de cohorte ou cas-témoins. La majorité concernait la population générale (81 études) ou des personnes atteintes de maladies spécifiques (19 études).

Concernant l'exposition à long terme (≥ 30 jours), 49 études ont été incluses, la plupart étant des cohortes. Pour l'association entre pollution à long terme et hypertension, il y avait 16 études pour les PM_{10} , 20 études pour les $\text{PM}_{2,5}$ et 19 études pour le NO_2 . Pour l'association entre pollution à long terme et pression artérielle systolique (PAS), il y avait 17 études sur les PM_{10} , 30 études sur les $\text{PM}_{2,5}$ et 9 études sur le NO_2 . Pour l'association entre pollution à long terme et pression artérielle diastolique (PAD), 9 études portaient sur les PM_{10} , 37 sur les $\text{PM}_{2,5}$ et PM_{10} sur le NO_2 .

La méta-analyse globale a montré une association entre les $\text{PM}_{2,5}$ (mais pas les PM_{10} ou le NO_2) et l'hypertension, une association entre les PM_{10} , les $\text{PM}_{2,5}$ et le NO_2 et la PAD, mais aucune association entre les polluants et la PAS. Les auteurs ont conclu à une hétérogénéité importante dans toutes les méta-analyses conduites.

Nous pouvons décrire quelques études incluses dans cette revue qui soulignent l'hétérogénéité dans les résultats. En effet, certaines études ont trouvé des associations positives entre l'exposition à la pollution atmosphérique et la pression sanguine. Par exemple, une étude conduite auprès de 12 605 adultes âgés de 50 ans et plus en Chine a montré une association significative entre l'exposition aux $\text{PM}_{2,5}$ et une augmentation de la PAS et de la PAD (109). Cependant, les associations ne sont pas toujours claires. En effet, par exemple, une étude transversale en Chine a trouvé une association significative entre l'exposition aux $\text{PM}_{2,5}$ et une augmentation de la PAS, mais aucune association avec la PAD (110). Une autre étude a examiné l'association entre l'exposition à la pollution atmosphérique liée au trafic et la pression sanguine chez 1 432 enfants issus de la cohorte PIAMA aux Pays Bas (pression sanguine mesurée quand les enfants étaient âgés de 12 ans) et a montré une augmentation significative de la PAD avec l'exposition. Mais cette étude n'a retrouvé aucune association significative avec la PAS (111). Enfin, une étude auprès de 2 612 personnes âgées à Bordeaux a montré une association négative entre l'exposition aux PM_{10} et la PAS mais cette étude n'a pas investigué l'association avec la PAD (112).

Depuis cette revue, d'autres études ont exploré le lien entre pollution et pression artérielle, notamment une étude en Chine sur un large échantillon de 14 622 adultes, montrant une augmentation de la pression artérielle avec une exposition à long terme à la pollution (113).

Globalement, la revue de littérature suggère une association entre pollution et pression artérielle, avec une minorité d'études en Europe de l'Ouest. La relation entre pollution et niveaux de lipides a davantage de résultats, avec des associations souvent retrouvées mais n'impactant pas toujours le même profil lipidique. Il est ainsi actuellement difficile d'émettre une conclusion définitive, en raison du nombre limité d'études qui étaient éligibles à l'inclusion dans la méta-analyse précédemment décrite (101). Par conséquent, malgré les

données existantes, de nouvelles études épidémiologiques restent nécessaires pour mieux clarifier ces liens.

L'objectif de la présente étude était d'explorer l'association entre les niveaux moyens d'exposition résidentielle aux NO₂, PM₁₀ et SO₂ d'une part, et les facteurs de risque cardiovasculaires — à savoir les paramètres lipidiques sanguins (TGs, CT, LDLs, HDLs) et la pression artérielle (PAS, PAD) — d'autre part.

2. Méthode

a) Population

Les données proviennent de l'enquête transversale ELISABET.

Pour la présente étude, nous avons sélectionné les participants d'ELISABET ayant résidé à la même adresse au cours des 12 mois précédents. Les participants présentant des données manquantes sur les polluants, l'âge, le sexe, la communauté urbaine (de Lille ou de Dunkerque), l'année d'inclusion, le niveau d'éducation, le statut tabagique, la consommation de fruits et légumes, le niveau d'activité physique, la consommation d'alcool, l'indice de masse corporelle et le statut diabétique ont été exclus. Nous avons retenu uniquement les participants disposant d'un profil lipidique complet : TG, LDL, HDL et CT. Par ailleurs, les participants sous traitement hypolipémiant ont été exclus, de même que ceux n'ayant pas été à jeun depuis au moins 10 heures au moment du prélèvement sanguin.

Pour les analyses portant sur le LDL, nous avons exclu les participants ayant un taux de TG supérieur à 4 g/L, afin de pouvoir calculer le LDL selon la formule de Friedewald ($LDL = CT - HDL - TG/5$ en g/L) (114).

b) Données sur les polluants

Comme décrit au-dessus, entre 2010 et 2013 à Lille et entre 2012 et 2013 à Dunkerque, les concentrations annuelles moyennes de NO₂, de PM₁₀ et (pour Dunkerque uniquement) de SO₂ ont été estimées par l'organisme de surveillance de la qualité de l'air ATMO Hauts-de-France (Lille, France) (48) à l'aide d'un système de modélisation de la dispersion atmosphérique (94).

Pour les 468 participants inclus en 2011 à Dunkerque, nous avons utilisé les niveaux de pollution de 2012 comme mesure d'exposition, car les données de concentration des polluants pour 2011 n'étaient pas disponibles (115).

c) Données socio-économiques sur les quartiers

La densité de population de chaque quartier a été calculée à partir des données extraites des bases de *Ilots Regroupés pour l'Information Statistique* (IRIS), unités administratives infra-communales gérées par *l'Institut national de la statistique et des études économiques* (INSEE : www.insee.fr) (47). L'unité IRIS correspond à la plus petite unité de recensement disponible en France. L'indice de défaveur sociale (FDep09, 2009), calculé au niveau de l'IRIS, a été utilisé pour tenir compte du niveau socioéconomique des participants.

d) Biomarqueurs lipidiques et pression artérielle

Les mesures ont déjà été décrites dans des études précédentes (47). Le cholestérol total (CT) et les triglycérides (TG) ont été mesurés à l'aide de méthodes enzymatiques standard. Le HDL a été mesuré après précipitation de l'apolipoprotéine B par le chlorure de phosphotungstate-magnésium. Comme mentionné précédemment, le taux de LDL a été calculé selon la formule

de Friedewald, et les participants étaient à jeun depuis au moins 10 h avant le prélèvement sanguin pour les dosages lipidiques.

De janvier à mars 2010, les analyses sanguines ont été réalisées au laboratoire d'analyse génomique de l'Institut Pasteur de Lille et au laboratoire central du Centre Hospitalier Universitaire de Lille. D'avril 2010 à novembre 2013, toutes les analyses sanguines ont été réalisées uniquement au laboratoire central.

Deux mesures consécutives de la pression artérielle systolique (PAS) et diastolique (PAD) ont été effectuées avec les participants en position assise, après 5 minutes de repos. Les examens ont été réalisés le matin, et les participants devaient être à jeun pour le rendez-vous. Lorsque cela était possible, la pression artérielle retenue correspondait à la moyenne de deux mesures. Une seconde mesure de la pression artérielle a été systématiquement effectuée uniquement durant le dernier tiers de la période d'étude (47,116).

L'hypertension a été définie par une PAS ≥ 140 mmHg, une PAD ≥ 90 mmHg, ou la prise d'un traitement antihypertenseur. Pour tenir compte de l'effet d'un traitement antihypertenseur dans nos analyses de pression artérielle, nous avons ajouté 10 mmHg et 5 mmHg respectivement aux valeurs observées de PAS et de PAD (47,117).

e) Analyses statistiques

Les associations entre l'exposition à chaque polluant atmosphérique (PM₁₀, SO₂ et NO₂) et les variables dépendantes (paramètres lipidiques sanguins et pression artérielle) ont été estimées à l'aide de modèles de régression linéaire multiple. Afin de normaliser la distribution des données, une transformation logarithmique a été réalisée sur les variables lipidiques sanguines. Ainsi, les coefficients de régression ont été exprimés en variation en pourcentage [intervalle de confiance (IC) à 95 %] par 2 $\mu\text{g}/\text{m}^3$ de PM₁₀ et de SO₂ ou par 5 $\mu\text{g}/\text{m}^3$ de NO₂. Ces incréments d'exposition résidentielle ont été choisis car ils étaient du même ordre de grandeur que l'intervalle interquartile des mesures à Lille et à Dunkerque. Les valeurs ont ensuite été rétro-transformées et converties en pourcentage de variation en multipliant par 100 puis en soustrayant 100 points de pourcentage. Les variables de pression artérielle n'ont pas nécessité de transformation logarithmique, leur distribution étant normale.

Pour chaque critère de jugement et chaque polluant, nous avons ajusté les modèles de régression linéaire généralisée sur des facteurs de confusion potentiels. Ainsi, le modèle 1 a été ajusté sur l'âge, le sexe, la communauté urbaine (de Lille ou de Dunkerque) et l'année d'inclusion (facteurs individuels de confusion). Pour le modèle 2, le niveau d'éducation, le statut tabagique, la consommation quotidienne de fruits et légumes, le niveau d'activité physique et la consommation d'alcool (facteurs potentiellement associés aux facteurs de risque cardiovasculaire) ont été ajoutés aux variables du modèle 1. Pour le modèle 3 complètement ajusté, l'indice de masse corporelle (IMC) et le statut diabétique (facteurs de médiation) ont été ajoutés aux variables du modèle 2.

Nous avons également testé les interactions entre chaque polluant et la communauté urbaine (de Lille ou de Dunkerque), l'âge, le sexe, l'IMC et le statut tabagique. De plus, nous avons formulé l'hypothèse selon laquelle les caractéristiques de la zone de résidence pouvaient être associées à la pollution et aux modifications du profil lipidique. Nous avons donc réalisé des analyses de sensibilité en ajoutant des facteurs géographiques (l'exposition définie selon le lieu de résidence) tels que l'indice de défaveur sociale (modèle 4) ou la densité de population (modèle 5).

Comme les modèles ont été définis a priori pour tenir compte des facteurs de confusion potentiels, nous avons conservé tous les facteurs potentiels dans le modèle et n'avons pas utilisé de critères d'ajustement (*goodness of fit*) pour sélectionner le meilleur modèle.

La validité des modèles de régression a été vérifiée en examinant la normalité des résidus de chaque modèle via l'histogramme des résidus, ce qui a permis de confirmer que ces derniers suivaient bien une distribution normale — condition nécessaire à la validité des modèles de régression.

Les variables d'ajustement ont été définies de la manière suivante :

- Âge : variable qualitative.
- Sexe : variable binaire (homme ou femme).
- Communauté urbaine : variable binaire (Lille ou Dunkerque).
- Année d'inclusion : variable quantitative en 3 classes (2011, 2012 et 2013).
- Niveau d'éducation : 4 classes (5 ans ou plus après le baccalauréat, 2 à 4 ans après le baccalauréat, pas d'études après le baccalauréat, ou pas d'études au lycée).
- Statut tabagique : 3 classes (n'a jamais fumé ; ancien fumeur ; fumeur actuel).
- Consommation de fruits et légumes : variable quantitative correspondant à la fréquence de consommation de fruits ajoutée à celle de légumes (nombre de portions par jour).
- Niveau d'activité physique : variable quantitative correspondant à la somme des équivalents métaboliques hebdomadaires pour activité intense ($8 \times$ nombre de jours \times durée moyenne), pour activité physique modérée ($4 \times$ nombre de jours \times durée moyenne) et pour la marche ($3,3 \times$ nombre de jours \times durée moyenne), selon le *Short International Physical Activity Questionnaire*.
- Consommation d'alcool : variable quantitative calculée comme suit : nombre de verres par jour \times nombre de jours de consommation par semaine $\times 10 \div 7$, représentant la consommation quotidienne moyenne en grammes/jour. Les boissons considérées étaient le vin, le cidre, la bière, l'apéritif et autres alcools, la variable correspondant à la somme des consommations quotidiennes moyennes (en g/jour) de toutes ces boissons.
- IMC : variable quantitative.
- Statut diabétique : variable binaire définie comme « oui » si un médecin a déjà diagnostiqué un diabète au participant, ou si le participant prenait un traitement antidiabétique, ou si la glycémie était $\geq 1,26$ g/L (à jeun), ou si l'HbA1c était $\geq 6,5$ %, ou si la glycémie était ≥ 2 g/L. Sinon, la variable prenait la valeur « non ».

Le seuil de significativité statistique a été fixé à $p < 0,05$. Les analyses statistiques ont été réalisées avec le logiciel R (version 4.3.0).

3. Résultats de cette étude

a) Population

L'enquête ELISABET comptait 3 276 participants. Parmi eux, 101 ont été exclus car ils avaient retiré leur consentement à l'utilisation de leurs données personnelles.

Pour les analyses portant sur les TG, le TC et le HDL, 43 participants ont été exclus (ramenant le total à 3 132) car ils avaient résidé moins de 12 mois à l'adresse déclarée. Ensuite, 214 participants ont été exclus en raison de données manquantes sur les covariables.

Parmi les 2 918 participants restants, 17 ont été exclus en raison de données manquantes sur les données lipidiques et 466 car ils suivaient un traitement hypolipémiant. Il restait donc 2 435 participants.

Après l'exclusion de 96 participants qui n'étaient pas à jeun avant le prélèvement sanguin, l'échantillon final comprenait 2 339 participants.

Pour l'analyse sur le LDL, nous avons exclus 18 participants ayant des données manquantes sur le LDL et 12 participants ayant un taux de TG supérieur à 4 g/L (rendant impossible l'utilisation de la formule de Friedewald) (25). L'analyse du LDL a donc porté sur 2 309 participants.

Concernant les analyses sur la pression artérielle, 42 participants présentant des données manquantes pour la PAS et la PAD ont été exclus, ce qui a abouti à un échantillon de 2297 individus.

Les caractéristiques des participants inclus dans l'analyse complète sont résumées dans le tableau III.A.3.1. L'âge moyen (écart-type) était de 52,26 ans (7,07) à Lille et de 52,26 ans (7,18) à Dunkerque. La proportion d'hommes était de 44,20 % (n = 554) à Lille et de 46,70 % (n = 507) à Dunkerque. Comparés aux participants de Lille, ceux de Dunkerque présentaient un IMC plus élevé et des prévalences plus importantes de diabète et de consommation d'alcool, mais un niveau d'éducation, une activité physique et une consommation quotidienne de fruits et légumes plus faibles.

b) Pollution

Les concentrations de PM₁₀ et de NO₂ étaient plus élevées à Lille qu'à Dunkerque (27,02 µg/m³ contre 26,53 µg/m³ pour les PM₁₀, et 26,05 µg/m³ contre 19,96 µg/m³ pour le NO₂, respectivement ; tableau III.A.3.1).

Le niveau moyen d'exposition au SO₂ (mesuré uniquement à Dunkerque) était de 3,28 µg/m³. Le coefficient de corrélation de Pearson entre le NO₂ et les PM₁₀ était de 0,87 à Lille et de 0,83 à Dunkerque.

Le coefficient était de 0,64 pour la corrélation entre le SO₂ et le NO₂, et de 0,72 pour la corrélation entre le SO₂ et le PM₁₀.

c) Associations entre pollution de l'air résidentielle et facteurs de risque cardiovasculaire

Les analyses multivariées ont montré une association significative entre l'exposition aux PM₁₀ et la concentration de HDL (tableau III.A.3.2). Ainsi, une augmentation de 2 µg/m³ de PM₁₀ était significativement associée à une diminution de 1,72 % du taux de HDL ($p = 0,0037$).

La concentration de HDL diminuait également de manière significative avec l'exposition au NO₂ et au SO₂, mais uniquement dans le modèle 1 (diminutions respectives de 1,14 % ($p = 0,0461$) et de 2,14 % ($p = 0,0097$)).

À Dunkerque, l'association entre l'exposition au SO₂ et le taux de TG observée dans le modèle 1 n'était plus significative dans le modèle complet (modèle 3 : $p = 0,162$).

Dans l'analyse multivariée, aucune association significative n'a été observée entre l'exposition à court terme ou résidentielle aux polluants atmosphériques et la pression artérielle diastolique ou systolique (à l'exception d'une association significative entre l'exposition au SO₂ et la PAS, mais uniquement dans le modèle 1) (tableau III.A.3.3).

d) Analyses de sensibilité

Dans les analyses de sensibilité (tableau III.A.3.4), l'association entre l'exposition aux PM₁₀ et le taux de HDL restait significative dans le modèle 4 (c'est-à-dire le modèle 3 ajusté sur un indice de défavorisation sociale ; $p = 0,023$).

Les interactions entre la zone urbaine, l'IMC, le statut tabagique, l'âge, le sexe et les niveaux de pollution n'étaient pas statistiquement significatives.

Tableau III.A.3.1. Caractéristiques des participants de l'étude

	Lille	Dunkerque
n (%)	1254 (53,61)	1085 (46,39)
Sexe (hommes), n (%)	554 (44,20)	507 (46,70)
Age (années), moyenne (ET)	52,26 (7,07)	52,26 (7,18)
Fumeurs, n (%)	242 (19,30)	196 (18,10)
IMC (kg/m ²), moyenne (ET)	26,11 (4,88)	26,94 (4,90)
Niveau d'éducation, n (%)		
5 ans ou plus après le baccalauréat	320 (25,50)	157 (14,50)
2 à 4 ans après le baccalauréat	280 (22,30)	176 (16,20)
pas d'études après le baccalauréat	570 (45,50)	640 (59,00)
pas d'études au lycée	84 (6,70)	112 (10,30)
Niveau d'activité physique (MET), moyenne (ET)	3464 (4528)	2058 (3041)
Diabète, n (%)	96 (7,70)	109 (10,00)
Consommation de fruits et légumes (par jour), moyenne (ET)	2,53 (1,34)	2,33 (1,30)
Consommation d'alcool (g/jour), moyenne (ET)	8,83 (13,48)	10,24 (15,77)
Profil lipidique (g/L), médiane [IIQ]		
HDL	0,58 [0,49 ; 0,69]	0,56 [0,48 ; 0,67]
Triglycérides	0,92 [0,70 ; 1,28]	0,96 [0,72 ; 1,35]
Cholestérol total	2,24 [1,98 ; 2,48]	2,24 [1,99 ; 2,50]
LDL	1,41 [1,21 ; 1,64]	1,44 [1,21 ; 1,66]
Pression artérielle (mmHg), moyenne (ET)		
PAS	125,84 (18,76)	128,93 (19,12)
PAD	82,02 (11,68)	82,57 (12,72)
Polluants (µg/m ³), moyenne (ET)		
SO ₂		3,28 (1,62)
PM ₁₀	27,02 (1,98)	26,53 (1,00)
NO ₂	26,05 (5,31)	19,96 (2,19)

Abréviations : ET = écart type ; IMC = indice de masse corporelle ; MET = équivalent métabolique ; IIQ = intervalle interquartile ; HDL = lipoprotéine de haute densité (high-density lipoprotein) ; LDL = lipoprotéine de basse densité (low-density lipoprotein) ; PAS = pression artérielle systolique ; PAD = pression artérielle diastolique ; SO₂ = dioxyde de soufre ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; NO₂ = dioxyde d'azote

Tableau III.A.3.2. Associations entre l'exposition aux polluants atmosphériques et les biomarqueurs sériques (g/L)

Biomarqueur	NO ₂			PM ₁₀			SO ₂		
	Effectif	Pourcentage de variation pour 5 µg/m ³ [IC 95%]	p	Effectif	Pourcentage de variation pour 2 µg/m ³ [IC 95%]	p	Effectif	Pourcentage de variation pour 2 µg/m ³ [IC 95%]	p
Triglycérides									
Modèle 1	2339	0,46 [-1,79 ; 2,77]	0,690	2339	2,36 [-0,17 ; 4,96]	0,068	1085	4,48 [1,03 ; 8,04]	0,0107
Modèle 2	2339	0,10 [-2,16 ; 2,41]	0,931	2339	1,62 [-0,91 ; 4,21]	0,213	1085	3,31 [-0,10 ; 6,84]	0,0573
Modèle 3	2339	0,28 [-1,87 ; 2,47]	0,801	2339	1,38 [-1,01 ; 3,84]	0,260	1085	2,30 [-0,90 ; 5,62]	0,162
Cholestérol total									
Modèle 1	2339	-0,46 [-1,27 ; 0,37]	0,279	2339	-0,41 [-1,31 ; 0,51]	0,383	1085	0,63 [-0,59 ; 1,85]	0,314
Modèle 2	2339	-0,28 [-1,11 ; 0,56]	0,517	2339	-0,23 [-1,15 ; 0,69]	0,623	1085	0,44 [-0,77 ; 1,68]	0,477
Modèle 3	2339	-0,24 [-1,07 ; 0,60]	0,572	2339	-0,20 [-1,12 ; 0,72]	0,665	1085	0,53 [-0,69 ; 1,76]	0,396
HDL									
Modèle 1	2339	-1,14 [-2,25 ; -0,020]	0,0461	2339	-2,35 [-3,56 ; -1,13]	< 0,001	1085	-2,14 [-3,73 ; -0,53]	0,0097
Modèle 2	2339	-0,76 [-1,87 ; 0,35]	0,179	2339	-1,81 [-3,01 ; -0,59]	0,0038	1085	-1,92 [-3,48 ; -0,33]	0,0182
Modèle 3	2339	-0,87 [-1,92 ; 0,18]	0,105	2339	-1,72 [-2,86 ; -0,56]	0,0037	1085	-1,45 [-2,94 ; 0,060]	0,0607
LDL									
Modèle 1	2309	-0,65 [-1,79 ; 0,50]	0,266	2309	-0,47 [-1,73 ; 0,80]	0,466	1068	1,22 [-0,46 ; 2,92]	0,155
Modèle 2	2309	-0,47 [-1,63 ; 0,70]	0,426	2309	-0,31 [-1,59 ; 0,99]	0,637	1068	1,05 [-0,65 ; 2,77]	0,229
Modèle 3	2309	-0,40 [-1,55 ; 0,77]	0,502	2309	-0,27 [-1,55 ; 1,02]	0,681	1068	1,10 [-0,59 ; 2,82]	0,205

Abréviations : NO₂ = dioxyde d'azote ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; SO₂ = dioxyde de soufre ; HDL = lipoprotéine de haute densité (high-density lipoprotein) ; LDL = lipoprotéine de basse densité (low-density lipoprotein)

Modèle 1 : ajustement sur l'âge, le sexe, la communauté urbaine et l'année d'inclusion

Modèle 2 : ajustement sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, le niveau d'éducation, le statut tabagique, la consommation de fruits et légumes, le niveau d'activité physique et la consommation d'alcool

Modèle 3 : ajustement sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, le niveau d'éducation, le statut tabagique, la consommation de fruits et légumes, le niveau d'activité physique, la consommation d'alcool, l'indice de masse corporelle et le statut diabétique

Tableau III.A.3.3. Associations entre l'exposition aux polluants atmosphériques et la pression artérielle (mmHg)

Marqueur	NO ₂			PM ₁₀			SO ₂		
	Effectif	Variation du coefficient pour 5 µg/m ³ [IC 95%]	p	Effectif	Variation du coefficient pour 2 µg/m ³ [IC 95%]	p	Effectif	Variation du coefficient pour 2 µg/m ³ [IC 95%]	p
PAS									
Modèle 1	2297	-0,59 [-1,48 ; 0,30]	0,19	2297	-0,32 [-1,30 ; 0,66]	0,52	1065	1,37 [0,01 ; 2,73]	0,048
Modèle 2	2297	-0,25 [-1,14 ; 0,64]	0,58	2297	-0,11 [-1,09 ; 0,87]	0,83	1065	0,89 [-0,47 ; 2,26]	0,20
Modèle 3	2297	-0,16 [-1,03 ; 0,70]	0,71	2297	-0,14 [-1,09 ; 0,81]	0,77	1065	0,64 [-0,67 ; 1,96]	0,34
PAD									
Modèle 1	2297	-0,07 [-0,67 ; 0,53]	0,82	2297	0,11 [-0,55 ; 0,77]	0,75	1065	0,80 [-0,15 ; 1,74]	0,10
Modèle 2	2297	0,14 [-0,46 ; 0,74]	0,64	2297	0,23 [-0,43 ; 0,89]	0,49	1065	0,44 [-0,51 ; 1,39]	0,36
Modèle 3	2297	0,21 [-0,37 ; 0,79]	0,48	2297	0,22 [-0,42 ; 0,86]	0,50	1065	0,29 [-0,63 ; 1,21]	0,54

Abréviations : NO₂ = dioxyde d'azote ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; SO₂ = dioxyde de soufre ; PAS = pression artérielle systolique ; PAD = pression artérielle diastolique

Modèle 1 : ajustement sur l'âge, le sexe, la communauté urbaine et l'année d'inclusion

Modèle 2 : ajustement sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, le niveau d'éducation, le statut tabagique, la consommation de fruits et légumes, le niveau d'activité physique et la consommation d'alcool

Modèle 3 : ajustement sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, le niveau d'éducation, le statut tabagique, la consommation de fruits et légumes, le niveau d'activité physique, la consommation d'alcool, l'indice de masse corporelle et le statut diabétique

Tableau III.A.3.4. Associations entre l'exposition aux polluants atmosphériques et les biomarqueurs sériques (g/L) - analyses de sensibilité

Biomarqueur	NO ₂			PM ₁₀			SO ₂		
	Effectif	Pourcentage de variation pour 5 µg/m ³ [IC 95%]	p	Effectif	Pourcentage de variation pour 2 µg/m ³ [IC95%]	p	Effectif	Pourcentage de variation pour 2 µg/m ³ [IC 95%]	p
Triglycérides									
Modèle 4	2339	0,21 [-1,94 ; 2,40]	0,849	2339	0,86 [-1,56 ; 3,33]	0,490	1085	1,60 [-1,91 ; 5,23]	0,376
Modèle 5	2339	0,45 [-1,94 ; 2,91]	0,713	2339	1,27 [-1,42 ; 4,03]	0,359	1085	1,81 [-1,77 ; 5,52]	0,326
Cholestérol total									
Modèle 4	2339	-0,24 [-1,07 ; 0,60]	0,577	2339	-0,19 [-1,12 ; 0,75]	0,696	1085	0,71 [-0,64 ; 2,07]	0,303
Modèle 5	2339	0,09 [-0,84 ; 1,02]	0,857	2339	0,18 [-0,85 ; 1,23]	0,729	1085	0,98 [-0,39 ; 2,37]	0,162
HDL									
Modèle 4	2339	-0,82 [-1,86 ; 0,23]	0,125	2339	-1,36 [-2,52 ; -0,19]	0,023	1085	-0,80 [-2,45 ; 0,88]	0,351
Modèle 5	2339	-0,12 [-1,28 ; 1,07]	0,847	2339	-0,70 [-1,99 ; 0,61]	0,293	1085	-0,36 [-2,05 ; 1,36]	0,680
LDL									
Modèle 4	2309	-0,40 [-1,56 ; 0,77]	0,498	2309	-0,32 [-1,61 ; 1,00]	0,636	1068	1,09 [-0,77 ; 3,00]	0,253
Modèle 5	2309	-0,22 [-1,51 ; 1,09]	0,741	2309	-0,08 [-1,52 ; 1,38]	0,910	1068	1,34 [-0,57 ; 3,28]	0,172

Abréviations : NO₂ = dioxyde d'azote ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; SO₂ = dioxyde de soufre ; HDL = lipoprotéine de haute densité (high-density lipoprotein) ; LDL = lipoprotéine de basse densité (low-density lipoprotein)

Modèle 4 : ajustement sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, le niveau d'éducation, le statut tabagique, la consommation de fruits et légumes, le niveau d'activité physique, la consommation d'alcool, l'indice de masse corporelle, le statut diabétique et l'indice de défavorisation sociale (FDep)

Modèle 5 : ajustement sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, le niveau d'éducation, le statut tabagique, la consommation de fruits et légumes, le niveau d'activité physique, la consommation d'alcool, l'indice de masse corporelle, le statut diabétique, l'indice de défavorisation sociale (FDep) et la densité de population

4. Discussion de cette étude

Nos résultats ont montré qu'une diminution du taux sanguin de HDL était associée à une exposition à long terme aux PM₁₀. Pour une augmentation de 2 µg/m³ de PM₁₀, nous avons observé une diminution de 1,72 % du taux de HDL dans le modèle 3 ($p = 0,0037$). Aucune association n'a été observée avec les autres facteurs de risque cardiovasculaire.

Dans la littérature, des associations sont rapportées entre l'exposition à long terme et les taux lipidiques, mais les biomarqueurs lipidiques impliqués ne sont pas toujours les mêmes. Dans une revue portant sur 22 études, 14 ont trouvé une association avec le TC, 13 avec le HDL, 9 avec le LDL et, enfin, 15 ont montré une association avec les TG. Les auteurs n'ont pu inclure que 3 études dans une méta-analyse et, bien que cette analyse ait suggéré une association entre l'exposition aux PM₁₀ et au NO₂ et les TG, la conclusion était que des preuves supplémentaires étaient nécessaires (101). Ces divergences restent à expliquer. Elles pourraient être liées aux caractéristiques des populations étudiées et aux expositions aux polluants (durée, niveau et source) et/ou à des différences dans les méthodes de mesure.

L'association entre la pollution atmosphérique et le taux de HDL a une explication biologique plausible. L'exposition aux polluants atmosphériques déclenche une cascade de signaux inflammatoires, notamment via la synthèse de la sérum amyloïde A (SAA). Cette dernière remplace l'apolipoprotéine (Apo) A1, qui constitue un composant de la lipoprotéine HDL et un cofacteur de la lécithine-cholestérol acyltransférase (enzyme qui estérifie le cholestérol avant son transport vers le foie). Une fois dans le foie, le cholestérol est en partie recyclé et éliminé via le cycle entéro-hépatique. De plus, le SAA-HDL facilite l'efflux du cholestérol vers les macrophages de l'intima vasculaire, ce qui favorise leur transformation en cellules spumeuses et accélère la formation de plaques d'athérome (118,119).

Une autre hypothèse serait que les particules polluantes se lient directement aux enzymes impliquées dans la lipolyse. Une diminution de l'activité de ces enzymes pourrait conduire à une augmentation des TG et à une diminution du HDL (120). Toutefois, contrairement à cette hypothèse, nous n'avons pas observé d'association significative avec les TG après ajustement. Cela pourrait s'expliquer par la plus grande variabilité des mesures de TG, qui pourrait réduire la puissance statistique.

La diminution de la force de l'association entre la pollution et le HDL après ajustement pour les caractéristiques des quartiers dans l'analyse de sensibilité suggère qu'une partie de cette association est due à des facteurs géographiques de confusion. Néanmoins, les caractéristiques des quartiers peuvent avoir une association causale directe ou indirecte avec la pollution atmosphérique. Cela est particulièrement vrai pour la densité de population. La diminution de la force de l'association pourrait donc également s'expliquer par un surajustement (*overfitting*).

Concernant l'association entre pollution et pression artérielle, nos résultats ne sont pas significatifs. Dans la littérature, plusieurs études ont montré des associations significatives entre l'exposition à la pollution et l'augmentation de la pression artérielle (108,121). Plusieurs mécanismes ont été proposés. Tout d'abord, les particules polluantes peuvent altérer les médiateurs de l'inflammation tels que les cytokines, ainsi que certaines molécules vasoactives comme l'endothéline, puissant vasoconstricteur (108). De plus, les particules polluantes peuvent perturber la régulation du système nerveux autonome, entraînant une augmentation du tonus sympathique et une diminution du tonus parasympathique, ce qui conduit à une

élévation de la pression artérielle (108,122). En outre, les particules polluantes peuvent pénétrer dans la circulation sanguine en traversant la membrane alvéolaire, et altérer l'endothélium vasculaire (108). Enfin, une autre hypothèse est que les particules fines pourraient perturber la méthylation de l'ADN, pouvant ainsi entraîner des modifications de la pression artérielle (108,123).

Dans notre étude, l'absence d'association statistiquement significative entre pollution et pression artérielle pourrait s'expliquer par un manque de puissance statistique. En effet, la réalisation d'une seule mesure de la pression artérielle chez la majorité des participants introduit un biais de mesure non différentiel (124). De plus, l'écart-type — c'est-à-dire la différence entre les sujets les plus exposés et les moins exposés — est faible pour les expositions à long terme.

Forces et limites

Notre étude présente certaines limites. Premièrement, l'exposition aux polluants atmosphériques a été modélisée uniquement à partir du domicile des participants ; ainsi, l'exposition professionnelle et les autres expositions lors d'activités extérieures n'ont pas été prises en compte. Deuxièmement, bien que nous ayons ajusté sur de nombreuses covariables, l'association entre le niveau de pollution atmosphérique et le taux de HDL pourrait avoir été biaisée par des facteurs de confusion non mesurés (par exemple, des marqueurs d'inflammation tels que la SAA). Troisièmement, nous avons mesuré les taux lipidiques à partir d'échantillons sanguins prélevés après un jeûne d'au moins 10 heures, au lieu des 12 heures recommandées par les directives. En outre, nous ne disposons que d'une seule mesure des biomarqueurs, ce qui peut entraîner un manque de puissance statistique. Enfin, nous n'avons pas corrigé pour les comparaisons multiples, car les tests statistiques n'étaient pas indépendants. Néanmoins, l'association entre le HDL et les PM₁₀ restait statistiquement significative après correction pour 12 tests (c'est-à-dire quatre marqueurs lipidiques avec trois polluants atmosphériques).

L'étude présente également plusieurs points forts. Premièrement, nous avons utilisé des méthodes et protocoles de mesure standardisés pour analyser un échantillon représentatif de la population générale. Deuxièmement, l'exposition aux polluants a été évaluée à l'aide d'un protocole couramment utilisé pour élaborer les indices de qualité de l'air. Troisièmement, l'exposition aux polluants au domicile des participants a été modélisée avec une résolution spatiale élevée.

5. Conclusion de cette étude

Les résultats de notre étude transversale menée auprès d'une population urbaine d'âge moyen ont mis en évidence une association entre l'exposition à long terme aux PM₁₀ et de faibles taux de HDL. Les associations avec la pression artérielle et les autres facteurs lipidiques n'étaient pas significatives. Cette observation suggère que la pollution atmosphérique ambiante constitue un facteur de risque de modifications du métabolisme du HDL. Des études longitudinales supplémentaires sont nécessaires pour explorer la nature causale potentielle de cette association.

B. Etudier les relations entre exposition et pathologie à l'aide d'un intermédiaire de type biomarqueur

L'article publié est disponible en **annexe 2**.

1. Introduction

La pollution de l'air a un impact majeur sur la santé : selon l'OMS, elle provoque chaque année la mort de 7 millions de personnes (125). Les effets de la pollution de l'air sur les systèmes respiratoire et cardiovasculaire sont particulièrement nocifs (97). L'inflammation systémique de bas grade n'est pas causée par une maladie chronique ou aiguë spécifique, mais elle favorise l'apparition et/ou la progression de certaines maladies chroniques, notamment les maladies cardiovasculaires (126), le diabète (127), le cancer (128) et même la dépression (129). Ainsi, une inflammation pulmonaire et systémique de bas grade pourrait augmenter le potentiel de coagulation, de thrombose et, en fin de compte, le risque d'événements cardiovasculaires et ischémiques (130).

De plus, les particules en suspension (PM) et certains de leurs constituants ont des effets pro-inflammatoires sur l'organisme. Les PM pourraient accroître l'inflammation de bas grade et ainsi contribuer aux problèmes de santé susmentionnés. Plus précisément, les modifications biologiques induites par l'exposition aux PM peuvent survenir via plusieurs voies mécanistiques (97,131) :

- (i) la libération de médiateurs pro-inflammatoires (tels que les cytokines) et d'espèces réactives de l'oxygène par les cellules pulmonaires ;
- (ii) la perturbation du système nerveux autonome par l'interaction entre les particules inhalées et les nerfs pulmonaires ; et
- (iii) le passage direct dans la circulation systémique de particules ultrafines ou de certains constituants de ces particules.

Les marqueurs biologiques de l'inflammation de bas grade incluent la protéine C-réactive ultrasensible (CRP_{us}) et les cytokines inflammatoires. La CRP_{us} est un marqueur de l'inflammation de bas grade fréquemment mesuré dans les études épidémiologiques, alors que les autres cytokines sont dosées moins souvent. La CRP est l'un des réactifs de phase aiguë les plus caractéristiques. Ses concentrations plasmatiques augmentent rapidement et de manière significative en réponse à une inflammation, une lésion cellulaire ou une atteinte tissulaire. La CRP est ainsi depuis longtemps considérée comme un marqueur de référence de l'inflammation. Les tests de CRP ultrasensible permettent de mesurer avec précision de faibles augmentations dans l'intervalle normal de la CRP, et peuvent ainsi détecter des niveaux faibles mais persistants d'inflammation. Étant donné que l'athérosclérose comporte une composante inflammatoire, le dosage de la CRP_{us} apporte des informations sur le risque cardiovasculaire (132).

Dans les études portant sur les mécanismes cellulaires et moléculaires sous-jacents de l'inflammation, les concentrations de cytokines sont généralement mesurées dans des échantillons de sérum ou de plasma, prélevés chez des patients ou des individus en bonne santé. Les cytokines sont en effet des protéines solubles qui agissent comme des messagers chimiques de la réponse immunitaire et jouent un rôle important dans la communication intercellulaire, la croissance cellulaire et la différenciation cellulaire. Par ailleurs, une sécrétion anormale de certaines cytokines a été associée à la pathogenèse ou à la progression de maladies inflammatoires chroniques, ainsi qu'à l'exposition à des contaminants chimiques ou biologiques présents dans l'environnement.

Dans les grandes études épidémiologiques comme dans la pratique clinique courante, la détection et la quantification des cytokines peuvent utilement compléter les données issues du dosage de la CRP. Dans la présente étude, nous avons mesuré sept cytokines (interleukines (IL) -1 β , -6, -8, -10, -17A et -22, ainsi que le facteur de nécrose tumorale alpha (TNF α)), connues pour être impliquées dans les maladies inflammatoires ou dans la réponse à l'exposition aux polluants atmosphériques (133–136).

L'association entre la pollution de l'air et l'inflammation a été étudiée dans des populations présentant une condition de santé spécifique ou une exposition particulière, sur des échantillons restreints (de 15 à 175 sujets), et ces études hétérogènes ont donné des résultats incohérents (131,137–156).

L'association entre la CRPus et l'inflammation a également été étudiée chez les femmes enceintes, montrant qu'une CRPus plus élevée était significativement associée à l'exposition annuelle moyenne aux PM₁₀, mais pas au dioxyde d'azote (NO₂) (157). Chez les nouveau-nés à terme, une association a été observée entre les PM₁₀ et l'IL-1 β (158). Chez les enfants asthmatiques et non asthmatiques, une association a été retrouvée entre le NO₂ et l'IL-6 et l'IL-10 (159).

Les données de la littérature restent cependant discordantes : certains chercheurs ont même observé des niveaux plus faibles de CRP (145,152,153) et de cytokines pro-inflammatoires (144,147,154,155,158), ou des niveaux plus élevés d'une cytokine anti-inflammatoire (146,159) associés à une exposition plus importante.

Seules quelques études ont évalué l'association entre les marqueurs inflammatoires et la pollution de l'air chez des adultes non enceintes, dans des études en population générale, laissant cette association entre pollution atmosphérique et inflammation en population générale insuffisamment quantifiée. La plupart de ces études ne mesuraient que les niveaux de CRPus, et les résultats étaient incohérents. Une association significative entre les PM, le NO₂ et la CRPus a été observée chez des adultes chinois âgés ; toutefois, le niveau de pollution atmosphérique y était très élevé, atteignant 91 $\mu\text{g}/\text{m}^3$ (160).

Certaines études menées dans des régions d'Europe ou d'Amérique, où les niveaux de pollution sont très modérés, n'ont pas mis en évidence d'association significative. Par exemple, une étude réalisée en Allemagne auprès de 2 252 personnes n'a trouvé aucune association significative entre les PM, les NOx ou l'ozone (O₃), d'une part, et la CRPus, d'autre part (152). De même, dans une étude menée auprès de 408 résidents de Boston (États-Unis), les PM n'étaient pas significativement associées à la CRPus (161). Dans un échantillon de 17 000 personnes au Royaume-Uni, aucune association entre les PM₁₀ et la CRPus n'a été observée (153). Une étude portant sur 3 860 membres de la population générale suisse n'a pas trouvé d'association significative entre les PM₁₀ et la CRP (162).

En revanche, d'autres études ont trouvé des associations significatives. Ainsi, une association positive statistiquement significative entre l'exposition annuelle aux PM_{2,5} et aux PM₁₀ et le niveau sérique de CRPus a été observée dans un échantillon de la population générale allemande, avec prélèvements réalisés lors de deux périodes de suivi (163). De plus, une association positive significative entre l'exposition annuelle aux PM_{2,5}, l'exposition sur 6 mois à l'O₃ et le niveau sérique de CRPus (avec mesures répétées) a été retrouvée dans une étude menée auprès de 2 086 femmes aux États-Unis (130). Enfin, dans une analyse transversale de l'étude Enquête Littoral Souffle Air Biologie Environnement (ELISABET) menée dans le nord de la France, nous avons précédemment observé une association entre le NO₂ mesuré le jour du prélèvement sanguin et la veille, et le niveau sérique de CRPus, dans un sous-groupe spécifique de 1 506 non-fumeurs en bonne santé (c'est-à-dire sans maladie respiratoire chronique ou

aiguë, sans prise de traitement pulmonaire, sans obstruction des voies aériennes, sans exposition au tabagisme actif ou passif, et sans syndrome inflammatoire — défini par une CRPus sérique > 10 mg/L) (7).

Concernant les cytokines, les données publiées sont plus limitées. Dans une étude portant sur 3 860 individus de la population générale suisse (incluant deux périodes de suivi), des associations positives significatives ont été observées entre l'exposition aux particules de diamètre aérodynamique < 10 µm (PM₁₀) sur 1 jour, 1 semaine, 1 mois, 3 mois et 6 mois, et les cytokines inflammatoires (IL-6, IL-1β et TNFα) (162). Toutefois, aucune association n'a été observée entre les PM₁₀ et la CRPus (162). Ces associations significatives avec les cytokines n'ont pas encore été confirmées dans d'autres études. Comme mentionné précédemment, l'étude menée auprès de 408 résidents de Boston (161) n'a montré aucune association significative. Par ailleurs, ces études en population adulte (161,162) ont mesuré un nombre limité de cytokines (IL-1β, IL-6 et TNFα).

L'objectif de la présente étude était d'explorer :

- (i) les associations entre les niveaux de polluants atmosphériques et les concentrations sériques de CRPus dans l'ensemble de la population de l'étude ELISABET, et
- (ii) les associations entre les niveaux de polluants atmosphériques et les concentrations plasmatiques d'un large panel de cytokines dans un sous-échantillon de non-fumeurs résidant dans la zone urbaine de Lille.

Nous avons examiné à la fois l'exposition à court terme (c'est-à-dire le jour du prélèvement sanguin et la veille) et l'exposition à long terme (c'est-à-dire l'exposition annuelle moyenne à l'adresse de résidence), et analysé les échantillons sanguins pour mesurer les biomarqueurs.

2. Méthode

a) Population

La population de l'étude comprenait des participants hommes et femmes (âgés de 40 à 65 ans) issus de l'enquête transversale ELISABET réalisée entre 2011 et 2013 dans le nord de la France. Tous les participants résidaient depuis au moins cinq ans dans la même ville (Lille ou Dunkerque) ou dans son agglomération, et avaient été recrutés entre janvier 2011 et novembre 2013 (91).

Les participants ont été sélectionnés à partir des listes électorales par échantillonnage aléatoire, avec stratification selon le sexe, l'âge et la communauté urbaine (de Lille ou de Dunkerque), puis contactés dans un ordre aléatoire. Les données ont été recueillies au domicile des participants ou, plus rarement, lors d'une consultation dans un établissement de santé. Dans tous les cas, un(e) infirmier(ère) diplômé(e) d'État, spécialement formée, administrait un questionnaire détaillé. Un prélèvement sanguin était réalisé au cours de la même visite (7).

Nous avons exclu de l'étude les participants présentant un syndrome inflammatoire connu (défini par une concentration sérique de CRPus > 10 mg/L), ainsi que ceux dont il manquait les données relatives à la CRPus, aux niveaux de polluants ou aux covariables. Nous avons ensuite étudié d'autres biomarqueurs inflammatoires (IL-1β, IL-6, IL-8, IL-10, IL-17A, IL-22 et TNFα) dans un sous-échantillon de 1 125 non-fumeurs, afin d'éviter le biais lié à l'inflammation induite par le tabagisme. Nous avons inclus uniquement les participants de la zone de Lille, car l'écart-type du niveau résidentiel des PM₁₀ y était plus élevé qu'à Dunkerque (1,936 µg/m³ contre 0,999 µg/m³, respectivement), ce qui augmentait ainsi la puissance statistique de l'analyse sur un effectif plus restreint. Afin de permettre de futures études génétiques, nous avons inclus des participants d'origine

caucasienne disposant de données sur les variables génétiques. Enfin, nous avons exclu les participants présentant un syndrome inflammatoire connu, des données de spirométrie inacceptables, n'étant pas à jeun ou des données manquantes.

L'étude a été enregistrée sur ClinicalTrials.gov (NCT02490553) et le protocole a été approuvé par le comité de protection des personnes Nord-Ouest IV (Lille, France ; référence : 2010-A00065-34), conformément à la législation française sur la recherche biomédicale. Tous les participants ont donné leur consentement éclairé, écrit et signé pour participer à l'étude (7).

b) Zones d'étude

Nous avons étudié des participants vivant dans deux zones urbaines du nord de la France (Lille et Dunkerque, incluant la ville principale et sa proche banlieue). Une carte situant ces villes est présentée en figure III.B.3.1.

Nous avons d'abord vérifié que les caractéristiques des deux zones urbaines répondaient aux critères APHEKOM, c'est-à-dire une exposition homogène à un polluant atmosphérique donné, un jour donné, sur l'ensemble de la zone d'étude (164). En substance, cela signifie que la majorité de la population devrait vivre et travailler à l'intérieur de la zone d'étude ; aucune partie de cette zone ne devrait être sous l'influence d'une source importante et stationnaire de pollution susceptible d'entraîner une forte hétérogénéité des niveaux de polluants ; la zone urbaine étudiée ne devrait pas être fragmentée par de vastes zones non urbaines ; et il ne devrait pas exister de variations topographiques majeures susceptibles de modifier la dispersion et les concentrations des polluants (7).

La zone de Lille est une zone urbaine dense, comportant seulement quelques espaces non urbains entre les trois principales villes (Lille, Roubaix et Tourcoing) et deux plus petites villes (Armentières et Comines). Le réseau routier et autoroutier y est très dense. La zone urbaine de Lille est traversée par cinq autoroutes ; en 2012, le nombre moyen de véhicules par jour variait de 36 000 à 176 000 (5,6).

La zone de Dunkerque est une zone urbaine côtière. La mer se trouve au nord de la ville, tandis qu'un port et une zone industrielle sont situés au nord-ouest. Les vents dominants soufflent du sud-ouest ; par conséquent, la pollution atmosphérique provenant de la zone industrielle est en grande partie poussée vers la mer. Le réseau routier et autoroutier de cette zone est également dense (7). La zone urbaine de Dunkerque est traversée par deux autoroutes ; en 2012, le nombre moyen de véhicules par jour variait de 27 000 à 61 000 (5,6).

c) Pollution atmosphérique et mesures météorologiques

L'estimation de l'exposition à la pollution atmosphérique a été décrite précédemment (7,44,48). Nous avons examiné à la fois l'exposition à court terme et l'exposition résidentielle, en utilisant deux sources de données distinctes.

L'exposition à court terme a été définie comme le niveau moyen de pollution dans chaque zone d'étude (Lille ou Dunkerque) le jour de la mesure et la veille. Cette variable renseigne sur les effets à court terme et réversibles des périodes de forte pollution. L'exposition résidentielle a été définie comme l'exposition annuelle moyenne à l'adresse de résidence. Cette variable renseigne sur les différences entre les personnes vivant dans des zones fortement polluées et celles vivant dans des zones moins polluées, c'est-à-dire sur les effets à long terme de la pollution atmosphérique.

En résumé, l'exposition à court terme a été évaluée à partir des mesures quotidiennes de PM₁₀ et de dioxyde d'azote (NO₂), ainsi que des mesures horaires d'ozone (O₃), réalisées par les

stations de surveillance de la qualité de l'air exploitées par l'association ATMO Hauts-de-France (Lille, France). Les données provenant des stations urbaines et suburbaines répondaient aux critères d'homogénéité APHEKOM et ont été incluses. Les niveaux quotidiens de pollution pour chaque zone urbaine ont été définis comme la concentration moyenne journalière en PM₁₀ et en NO₂, et comme la moyenne mobile maximale sur 8 heures pour la concentration d'O₃, mesurées dans les stations de surveillance réparties dans les zones urbaines et suburbaines. Nous avons vérifié que les mesures quotidiennes provenant des stations situées dans une même zone urbaine étaient corrélées (7,164).

Les données météorologiques (y compris les précipitations quotidiennes, la température de l'air, l'humidité relative et la pression atmosphérique au niveau de la mer) ont été obtenues auprès de Météo-France.

Pour l'exposition résidentielle, les concentrations annuelles moyennes de PM₁₀ et de NO₂ entre 2010 et 2013 à Lille, et entre 2012 et 2013 à Dunkerque, ont été extraites des estimations produites par ATMO – Nord Pas de Calais, à l'aide du Atmospheric Dispersion Modelling System (ADMS, version 3.4 pour les zones urbaines), développé par Cambridge Environmental Research Consultants (165,166).

Ce modèle ne permettait pas d'estimer l'exposition résidentielle à l'O₃. Celui-ci intégrait des données météorologiques, topographiques et d'occupation des sols, les émissions de polluants d'origine naturelle et anthropique, ainsi que les données de pollution atmosphérique issues des stations de surveillance. Des cartes de concentration pour le PM₁₀ et le NO₂ à Lille et Dunkerque en 2015 sont présentées sur le site d'ATMO (ATMO Nord-Pas-de-Calais, 2016a, 2016b). Les cartes de distribution spatiale des polluants atmosphériques pour 2015 dans les zones urbaines de Lille et de Dunkerque sont disponibles sur le site d'ATMO Hauts-de-France (www.atmo-hdf.fr) (92,93).

Nous avons appliqué une résolution spatiale de 25 × 25 mètres. Le lieu de résidence de chaque participant a été localisé dans une maille de 25 mètres. Les niveaux d'exposition annuels au lieu de résidence ont été évalués par la valeur moyenne des quatre points les plus proches dans la grille, pondérée par l'inverse du carré de la distance à chacun de ces points (48,167).

d) Biomarqueurs

La concentration sérique de CRPus a été mesurée par un test néphélométrique (BN ProSpec System, Siemens) avec une plage de détection comprise entre 0,17 et 10 mg/L. Les valeurs inférieures à la limite de détection (LOD) ont été calculées comme 0,12 mg/L (LOD/√2) (7).

Pour la mesure des cytokines, nous avons utilisé le test immunologique multiplex à électrochimiluminescence MesoScale Discovery® (*Meso Scale Diagnostics LLC*, Rockville, MA, USA). Les tests multiplex (pour IL-1β, IL-6, IL-8, IL-10, IL-17A et TNFα) et le test individuel (pour IL-22) étaient fournis sur des plaques à 10 points (*U-PLEX Plate K15067L-2*, Meso Scale Diagnostics LLC). Les tests ont été réalisés conformément aux instructions du fabricant, à l'exception de la plage des courbes d'étalonnage, que nous avons modifiée. Le fabricant recommande des courbes d'étalonnage à huit points ; cependant, afin d'améliorer la sensibilité du test, nous avons ajouté deux points dans la partie basse de la courbe et retiré deux points de la partie haute.

Pour une plaque unique, 200 µL d'anticorps ont d'abord été incubés avec 300 µL de *linkers* pendant 30 minutes. Ensuite, 200 µL d'une solution d'arrêt ont été ajoutés, et le mélange a été incubé pendant 30 minutes, transféré dans des microplaques *QuickPlex 96-Well* (Meso Scale Diagnostics LLC) et incubé pendant 1 heure à température ambiante. Par la suite, 25 µL de diluant et 25 µL d'échantillons, d'étalons ou de contrôles ont été ajoutés en double, puis

incubés toute la nuit à 4 °C. Chaque puits a ensuite été lavé, incubé pendant 1 heure avec un anticorps de détection, lavé à nouveau, puis complété avec le *Read Buffer* (Meso Scale Diagnostics LLC). Les signaux d'électrochimiluminescence ont été détectés à l'aide de l'instrument QuickPlex SQ 120 et analysés avec le logiciel Discovery Workbench 4.0 (Meso Scale Diagnostics LLC).

Les valeurs inférieures à la limite basse de quantification (plus faible concentration de la courbe d'étalonnage présentant un coefficient de variation en pourcentage < 20 %) ont été considérées comme des concentrations réelles.

La stabilité du dosage des biomarqueurs au cours du temps a été vérifiée en analysant régulièrement des échantillons témoins pendant toute la durée de l'étude. À cet effet, nous avons constitué un pool de 40 échantillons plasmatiques résiduels prélevés lors de soins cliniques courants chez des patients adultes hospitalisés au CHU de Lille (Lille, France). Au moins un échantillon témoin était inclus dans chaque série d'analyses.

e) Analyses statistiques

Nous avons examiné à la fois la pollution à court terme et la pollution résidentielle. Nous avons calculé le coefficient de corrélation de Pearson pour les expositions résidentielles, pour chaque année et chaque ville. Les données d'exposition résidentielle pour l'année 2011 dans la zone de Dunkerque n'étaient pas disponibles (48). Nous avons imputé ces données de 2011 à partir des données d'exposition résidentielle de 2012 et de la moyenne des mesures quotidiennes issues des stations de surveillance de la pollution, en utilisant la formule suivante :

« Exposition résidentielle imputée en 2011 » = « exposition résidentielle en 2012 » – « moyenne annuelle des mesures quotidiennes en station de 2012 » + « moyenne annuelle des mesures quotidiennes en station de 2011 »

Afin d'évaluer la précision de cette imputation, nous avons imputé l'exposition résidentielle en 2012 à partir des données de 2013 et calculé l'accord entre les valeurs imputées et les valeurs disponibles à l'aide de la corrélation intraclasse sur l'ensemble de l'échantillon de l'étude ELISABET.

Les associations entre l'exposition à chaque polluant atmosphérique et les variables explicatives (niveaux de marqueurs inflammatoires) ont été estimées à l'aide de modèles de régression linéaire multiple. Afin de normaliser la distribution des données, toutes les variables explicatives ont été transformées en logarithmes. Ainsi, les coefficients de régression ont été exprimés comme le pourcentage de variation [intervalle de confiance (IC) à 95 %] par incrément de 10 µg/m³ de chaque polluant pour l'exposition à court terme, et par incrément de 2 µg/m³ de PM₁₀ ou 5 µg/m³ de NO₂ pour l'exposition résidentielle. Ces incréments pour l'exposition résidentielle ont été choisis car ils étaient du même ordre de grandeur que l'intervalle interquartile des mesures à Lille et à Dunkerque. Les valeurs ont ensuite été rétro-transformées et converties en pourcentage de variation en les multipliant par 100 puis en soustrayant 100 points de pourcentage.

Nous avons testé les interactions entre, d'une part, les niveaux de polluants et, d'autre part, la zone urbaine, le statut tabagique et l'IMC.

Pour l'exposition à court terme, les modèles ont été ajustés sur la zone urbaine (Lille ou Dunkerque), l'année d'inclusion, l'âge, le sexe, l'IMC, le statut tabagique (non-fumeurs / fumeurs actuels), le niveau d'éducation (5 ans ou plus après le baccalauréat, 2 à 4 ans après le baccalauréat, pas d'études après le baccalauréat, ou pas d'études au lycée), la densité de population (calculée à partir des bases de données IRIS des unités statistiques infracommunales obtenues auprès de l'Institut national de la statistique et des études

économiques – INSEE : www.insee.fr), les mesures météorologiques (humidité, pression atmosphérique, température de l'air et précipitations), la saison (hiver, printemps, été ou automne), le jour de la semaine, le nombre de jours depuis le début de l'étude et les périodes de vacances scolaires.

Pour l'exposition résidentielle, les modèles ont été ajustés sur la zone urbaine (Lille ou Dunkerque), l'année d'inclusion, l'âge, le sexe, l'IMC, le statut tabagique (non-fumeurs / fumeurs actuels) et le niveau d'éducation (5 ans ou plus après le baccalauréat, 2 à 4 ans après le baccalauréat, pas d'études après le baccalauréat, ou pas d'études au lycée).

Notre analyse principale portait sur les associations entre les variables explicatives et l'exposition à court terme aux PM₁₀, NO₂ et O₃, ainsi que sur les associations entre les variables explicatives et l'exposition résidentielle aux PM₁₀ et NO₂.

Nous avons ensuite réalisé deux analyses de sensibilité ; la première incluant les PM₁₀, le NO₂ et l'O₃ dans un modèle multipolluants ; et la seconde incluant la direction et la vitesse du vent ainsi que le cumul des heures d'ensoleillement de la journée (pour déterminer si le vent pouvait transporter la pollution industrielle vers la zone de Dunkerque).

Les données de vent et d'ensoleillement ont été obtenues auprès de la station météorologique régionale (Météo-France, Lille, France). L'ensoleillement a été inclus dans le modèle car la lumière solaire intervient dans la formation de l'ozone troposphérique via la photodissociation du NO₂ (168).

La direction du vent (moyenne horaire) a été mesurée à une hauteur de 10 m.

Nous avons comparé les jours avec un vent de secteur nord (compris entre 270° et 89° sur la boussole) aux jours avec un vent de secteur sud (compris entre 90° et 269° sur la boussole). La vitesse moyenne du vent (en m/s) a été calculée par heure, puis par jour. Nous avons exclu les vitesses horaires inférieures à 2 m/s, et calculé le nombre d'heures par jour avec un vent de secteur nord et le nombre d'heures par jour avec un vent de secteur sud. Nous n'avons pris en compte que les journées complètes avec un vent de secteur nord ou de secteur sud pendant au moins 70 % de la journée.

L'ensoleillement a été considéré comme la durée cumulée par jour. Cette variable n'étant pas mesurée à Dunkerque, nous avons utilisé les données de la station la plus proche (Calais-Marck, située à environ 30 km).

Toutes les analyses statistiques ont été réalisées avec le logiciel R (version 1.3.1093, *R Core Team*, *R Foundation for Statistical Computing*, <http://www.r-project.org>). Le seuil de significativité statistique a été fixé à $p < 0,05$.

3. Résultats de cette étude

a) Population

Un total de 3 275 participants a été inclus dans l'étude ELISABET.

Pour l'analyse sur la CRPus, nous avons exclu 198 participants (119 présentaient des données manquantes sur la CRPus et 79 un syndrome inflammatoire), ce qui a laissé 3 077 participants. Les analyses portant sur la pollution à court terme ont inclus 3 011 participants après exclusion de 26 participants avec des données manquantes sur au moins une covariable et de 40 participants avec des données manquantes pour les polluants (PM₁₀, NO₂ ou O₃). Les analyses sur la pollution résidentielle ont inclus 3 074 participants après exclusion de 2 participants ayant des données manquantes sur au moins une covariable manquante et d'un participant ayant des données manquantes pour les polluants (PM₁₀ ou NO₂).

Parmi les 3 275 participants, 1 125 ont été inclus dans l'analyse des cytokines. Dans la zone de Lille, on comptait 1 668 participants, dont 1 335 non-fumeurs (jamais fumeurs ou ayant arrêté

de fumer depuis au moins 12 mois). Des données génétiques étaient disponibles pour 1 125 de ces 1 335 participants. Nous avons ensuite exclu 72 participants avec des données de spirométrie non valides et 26 participants qui n'étaient pas à jeun au moment du prélèvement sanguin. Vingt-trois échantillons n'ont pas pu être analysés. Après l'exclusion de 22 participants présentant un syndrome inflammatoire, l'échantillon final comptait 982 participants.

Les caractéristiques des participants sont résumées dans le tableau III.B.3.1. Les populations de Lille et de Dunkerque ne différaient pas sur les principales caractéristiques, sauf pour l'IMC, plus élevé dans la population de Dunkerque, et le niveau d'éducation, plus élevé dans la population de Lille. Les caractéristiques des participants en fonction de la catégorie d'IMC (poids normal, surpoids ou obésité) sont présentées dans le tableau supplémentaire III.B.3.1.

b) Pollution

Les concentrations de polluants sont résumées dans le tableau III.B.3.1.

Pour l'exposition à court terme, les concentrations de PM₁₀ et de NO₂ étaient plus élevées à Lille qu'à Dunkerque, tandis que les concentrations d'O₃ étaient similaires dans les deux zones urbaines.

Pour l'exposition résidentielle, les concentrations de PM₁₀ et de NO₂ étaient significativement plus élevées à Lille qu'à Dunkerque.

Nous avons examiné la corrélation entre les différents polluants. Pour l'exposition à court terme, les PM₁₀ et le NO₂ étaient positivement corrélés entre eux ($r = 0,749$ à Dunkerque et $r = 0,635$ à Lille), tandis que l'O₃ était négativement corrélé avec le NO₂ ($r = -0,539$ à Dunkerque et $r = -0,660$ à Lille) et avec les PM₁₀ ($r = -0,261$ à Dunkerque et $r = -0,265$ à Lille). Pour l'exposition résidentielle, le coefficient de corrélation entre les PM₁₀ et le NO₂ était de 0,762 à Dunkerque et de 0,840 à Lille.

Pour l'exposition résidentielle, les coefficients de corrélation de Pearson entre les expositions mesurées en 2012 et 2013 à Dunkerque étaient de 0,94 pour les PM₁₀ et de 0,97 pour le NO₂. Pour les données de 2012 imputées à partir des données de pollution de 2013, la mesure de concordance par corrélation intraclasse était de 0,94 pour les PM₁₀ et de 0,89 pour le NO₂. Le niveau moyen d'exposition à Dunkerque était de 25,02 µg/m³ pour les PM₁₀ et 22,43 µg/m³ pour le NO₂ en 2012 ; et de 25,93 µg/m³ pour les PM₁₀ et 23,08 µg/m³ pour le NO₂ en 2013. La matrice de corrélation pour Lille est présentée dans le tableau supplémentaire III.B.3.2 ; les coefficients de corrélation y étaient très élevés.

c) Associations entre la pollution et les biomarqueurs inflammatoires

(1) Association entre la pollution et la CRPus :

Pour l'exposition à court terme (tableau III.B.3.2), un incrément de 10 µg/m³ de PM₁₀ était associé à une augmentation significative [IC 95 %] de la CRPus (+3,43 % [0,68 ; 6,25]) dans l'ensemble de la population étudiée ($p = 0,014$).

Un incrément de 10 µg/m³ de NO₂ était associé à une augmentation significative de la CRPus chez les résidents de Lille (+7,85 % [2,04 ; 13,98] ; $p = 0,0076$) mais pas chez les résidents de Dunkerque (-2,61 % [-7,55 ; 2,59] ; $p = 0,32$), avec une interaction significative (p interaction = 0,030).

Par ailleurs, l'association était significative chez les fumeurs actuels (+12,37 % [1,99 ; 23,8] ; $p = 0,019$) mais pas chez les non-fumeurs (-0,43 % [-4,35 ; 3,64] ; $p = 0,83$), avec une interaction statistiquement significative ($p = 0,018$).

Concernant l'O₃, l'association était statistiquement significative chez les résidents de Lille, avec une diminution de la CRPus (-5,32 % [-8,98 ; -1,50] ; $p = 0,0068$), mais pas chez les résidents de Dunkerque (+3,59 % [-0,72 ; 8,09] ; $p = 0,10$), et l'interaction était significative ($p = 0,025$). Nous avons également observé une interaction significative entre l'IMC et la pollution par les PM₁₀ ($p < 0,01$) et entre l'IMC et le NO₂ ($p = 0,043$). Le niveau de CRPus augmentait avec l'exposition aux PM₁₀ chez les participants ayant un IMC normal et chez ceux en surpoids, bien que l'association ne soit statistiquement significative que dans la classe de surpoids ($p = 0,0032$). Aucune association significative n'a été observée dans la classe obèse.

Pour l'exposition au NO₂, les variations de CRPus étaient similaires, mais les associations n'étaient pas significatives dans les classes d'IMC normal et surpoids ; en revanche, une diminution significative a été observée dans la classe obèse ($p = 0,029$). Après exclusion des participants obèses de l'analyse, les associations devenaient significatives : +6,17 % [2,72 ; 9,73] ($p < 0,001$) pour les PM₁₀ (p interaction $< 0,01$) et +5,83 % [1,06 ; 10,82] ($p = 0,016$) pour le NO₂ (p interaction obèses vs non-obèses = 0,016).

Par ailleurs, l'interaction avec la zone urbaine n'était plus significative après ajustement sur l'interaction entre pollution et IMC.

Pour l'exposition résidentielle (tableau III.B.3.2), les associations entre les polluants atmosphériques et le niveau de CRPus n'étaient pas statistiquement significatives. De même, les interactions entre l'exposition résidentielle aux polluants et l'IMC n'étaient pas statistiquement significatives.

(2) Association entre la pollution et les cytokines :

Aucune des associations entre l'IL-6, l'IL-8, l'IL-10, l'IL-22, le TNF α , l'IL-1 et l'IL-17A et l'exposition quotidienne (à court terme) aux PM₁₀ et au NO₂ n'était significative (tableau III.B.3.3).

Pour l'exposition à court terme à l'O₃ uniquement, l'association avec le TNF α était significative (+2,69 % [0,44 ; 4,99] ; $p = 0,019$).

Aucune des associations avec l'exposition résidentielle annuelle moyenne aux PM₁₀ ou au NO₂ n'était significative.

d) Analyses de sensibilité

Nous avons réalisé des régressions linéaires dans des modèles multi polluants.

Après ajustement sur l'O₃, les résultats pour l'exposition à court terme aux PM₁₀ (tableau supplémentaire III.B.3.3) suivaient la même tendance que les résultats principaux. Après ajustement sur le NO₂, l'association avec les PM₁₀ n'était plus significative chez les fumeurs actuels (+6,48 % [-2,81 ; 16,66] ; $p = 0,18$) mais devenait significative chez les non-fumeurs (+5,03 % [0,77 ; 9,47] ; $p = 0,020$), avec une interaction non significative ($p = 0,088$).

Pour l'exposition au NO₂, les associations n'étaient plus significatives après ajustement sur les PM₁₀, mais l'association avec la CRPus restait significative chez les fumeurs actuels après ajustement sur l'O₃ (+15,37 % [2,97 ; 29,26] ; $p = 0,014$), avec une interaction significative ($p = 0,017$).

Concernant l'exposition à l'O₃, les associations avec les niveaux de CRPus n'étaient pas significatives après ajustement sur les PM₁₀ ou le NO₂. L'interaction avec la zone urbaine demeurait significative.

Pour l'exposition résidentielle, les modèles multi polluants ne montraient aucune association statistiquement significative avec la CRPus (tableau supplémentaire III.B.3.4).

Dans les modèles multi polluants, les associations entre la pollution atmosphérique et les cytokines n'étaient pas statistiquement significatives.

Nous avons également effectué une régression linéaire en ajoutant les covariables météorologiques (direction du vent, vitesse du vent et ensoleillement) au modèle pour l'exposition à court terme et les niveaux de CRPus. Pour l'exposition aux PM₁₀, les associations avec la CRPus étaient statistiquement significatives dans l'ensemble de la population, chez les résidents de Lille et chez les non-fumeurs. Les associations entre le NO₂ ou l'O₃ et la CRPus n'étaient pas significatives. Les interactions entre polluants et zone urbaine, ainsi qu'entre polluants et statut tabagique, n'étaient plus significatives. Enfin, les interactions entre polluants et direction du vent n'étaient significatives ni à Lille ni à Dunkerque.

Tableau III.B.3.1. Caractéristiques des participants de l'étude

	Population CRPus N = 3074		Population Cytokines N = 982
	Dunkerque	Lille	
n (%)	1508 (49,06)	1566 (50,94)	982 (100)
Age (années), moyenne (ET)	53,26 (7,26)	53,16 (7,20)	53,84 (7,17)
Sexe (hommes), n (%)	735 (48,74)	742 (47,38)	451 (45,93)
Non-fumeurs, n (%)	1229 (81,50)	1279 (81,67)	982 (100)
IMC (kg/m ²), médiane (ET)	27,38 (5,08)	26,54 (4,96)	26,72 (5,03)
Niveau d'éducation, n (%)			
5 ans ou plus après le baccalauréat	193 (12,80)	376 (24,01)	263 (26,78)
2 à 4 ans après le baccalauréat	222 (14,72)	340 (21,71)	222 (22,61)
pas d'études après le baccalauréat	902 (59,81)	716 (45,72)	422 (42,97)
pas d'études au lycée	191 (12,67)	134 (8,56)	75 (7,64)
CRPus (mg/L), médiane [IIQ]	1,14 [0,58-2,47]	0,99 [0,50-2,04]	0,94 [0,49-1,89]
Biomarqueurs (pg/mL), médiane [IIQ]			
IL1 β			0,13 [0,13-0,13]
IL6			0,44 [0,30-0,67]
IL8			2,72 [2,09-3,76]
IL10			0,19 [0,13-0,28]
IL17a			1,39 [1,39-1,39]
TNF α			2,35 [1,92-2,90]
IL22			0,23 [0,13-0,42]
Polluants			
PM ₁₀ court terme ($\mu\text{g}/\text{m}^3$), médiane [IIQ]	21,50 [17,23-33,29]	22,50 [16,50-35,82]	22,56 [16,50-35,82]
NO ₂ court terme ($\mu\text{g}/\text{m}^3$), médiane [IIQ]	20,25 [13,75-29,50]	24,00 [18,00-32,72]	24,00 [18,12-32,44]
O ₃ court terme ($\mu\text{g}/\text{m}^3$), médiane [IIQ]	44,96 [32,84-55,06]	44,50 [30,97-56,51]	44,90 [31,16-56,43]
PM ₁₀ résidentielle ($\mu\text{g}/\text{m}^3$), médiane [IIQ]	26,54 [25,61-27,53]	27,03 [25,70-28,28]	26,87 [25,44-28,13]
NO ₂ résidentielle ($\mu\text{g}/\text{m}^3$), médiane [IIQ]	22,12 [19,68-26,10]	26,01 [22,62-28,63]	25,53 [22,15-28,28]

Abréviations : ET = écart type ; IIQ = intervalle interquartile ; IMC = indice de masse corporelle ; CRPus = protéine C-réactive ultrasensible ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 μm ; NO₂ = dioxyde d'azote ; O₃ = ozone

Tableau III.B.3.2. Associations entre l'exposition aux polluants atmosphériques et le taux sérique de CRPus (mg/L)

Population	Effectif	PM ₁₀			NO ₂			O ₃		
		Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction
Totale	3011	3,43 [0,68 ; 6,25]	0,014		1,75 [-1,96 ; 5,61]	0,36		-1,20 [-3,95 ; 1,64]	0,40	
Communauté urbaine				0,074			0,030			0,025
Lille	1554	5,83 [2,26 ; 9,54]	0,0013		7,85 [2,04 ; 13,98]	0,0076		-5,32 [-8,98 ; -1,50]	0,0068	
Dunkerque	1457	0,16 [-4,21 ; 4,73]	0,94		-2,61 [-7,55 ; 2,59]	0,32		3,59 [-0,72 ; 8,09]	0,10	
Tabac				0,082			0,018			0,84
Non-fumeurs	2461	2,22 [-0,71 ; 5,24]	0,14		-0,43 [-4,35 ; 3,64]	0,83		-1,01 [-3,99 ; 2,05]	0,51	
Fumeurs actuels	550	9,41 [2,16 ; 17,17]	0,010		12,37 [1,99 ; 23,80]	0,019		-1,82 [-8,78 ; 5,68]	0,63	
IMC				0,0074			0,043			0,57
IMC < 25	1163	3,15 [-1,74 ; 8,29]	0,21		3,34 [-3,38 ; 10,52]	0,34		-2,74 [-7,57 ; 2,34]	0,28	
25 ≤ IMC < 30	1167	6,69 [2,20 ; 11,37]	0,0032		4,26 [-1,86 ; 10,76]	0,18		-0,62 [-5,05 ; 4,01]	0,79	
IMC ≥ 30	681	-4,18 [-9,02 ; 0,92]	0,11		-7,47 [-13,66 ; -0,83]	0,029		2,79 [-2,59 ; 8,47]	0,32	
<i>exposition résidentielle</i>										
Population	Effectif	Pourcentage de variation pour 2 µg/m ³ [IC 95%]	p	p d'interaction	Pourcentage de variation pour 5 µg/m ³ [IC 95%]	p	p d'interaction			
Totale	3074	1,98 [-2,35 ; 6,51]	0,38		0,88 [-3,19 ; 5,13]	0,68				
Communauté urbaine				0,27			0,12			
Lille	1566	0,13 [-5,00 ; 5,55]	0,96		-0,91 [-5,27 ; 3,66]	0,69				
Dunkerque	1508	7,27 [-3,17 ; 18,83]	0,18		9,50 [-1,66 ; 21,92]	0,098				
Tabac				0,18			0,082			
Non-fumeurs	2508	3,24 [-1,42 ; 8,13]	0,18		2,84 [-1,56 ; 7,43]	0,21				
Fumeurs actuels	566	-3,44 [-14,46 ; 9,00]	0,57		-8,92 [-18,98 ; 2,38]	0,12				
IMC				0,28			0,18			
IMC < 25	1187	-0,23 [-7,31 ; 7,40]	0,95		-1,25 [-7,69 ; 5,65]	0,72				
25 ≤ IMC < 30	1190	2,93 [-3,83 ; 10,17]	0,40		3,76 [-3,03 ; 11,03]	0,29				
IMC ≥ 30	697	5,42 [-3,14 ; 14,72]	0,22		2,38 [-5,40 ; 10,80]	0,56				

Abréviations : CRPus = protéine C-réactive ultrasensible ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; NO₂ = dioxyde d'azote ; O₃ = ozone ; IMC = indice de masse corporelle (kg/m²)

Tableau III.B.3.3. Associations entre l'exposition à la pollution atmosphérique et les taux plasmatiques de biomarqueurs (pg/mL)

Biomarqueur	PM ₁₀		NO ₂		O ₃	
	<i>exposition à court terme</i>					
	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p
IL6	1,56 [-1,44 ; 4,65]	0,31	0,67 [-4,00 ; 5,57]	0,78	1,04 [-2,31 ; 4,52]	0,55
IL8	-1,39 [-3,64 ; 0,91]	0,23	-1,84 [-5,37 ; 1,83]	0,32	1,05 [-1,56 ; 3,72]	0,43
IL10	1,87 [-1,32 ; 5,16]	0,25	0,27 [-4,67 ; 5,47]	0,92	0,16 [-3,39 ; 3,83]	0,93
IL22	1,43 [-4,06 ; 7,24]	0,62	1,76 [-6,85 ; 11,16]	0,70	2,61 [-3,65 ; 9,27]	0,42
TNF	-0,72 [-2,65 ; 1,25]	0,47	-2,25 [-5,25 ; 0,83]	0,15	2,69 [0,44 ; 4,99]	0,019
IL1	-1,81 [-6,68 ; 3,32]	0,48	-1,25 [-8,96 ; 7,10]	0,76	1,07 [-4,63 ; 7,11]	0,72
IL17a	-0,23 [-3,95 ; 3,63]	0,91	-4,20 [-9,76 ; 1,72]	0,16	3,19 [-1,17 ; 7,75]	0,15
<i>exposition résidentielle</i>						
Biomarqueur	Pourcentage de variation pour 2 µg/m ³ [IC 95%]	p	Pourcentage de variation pour 5 µg/m ³ [IC 95%]	p		
IL6	-0,39 [-4,68 ; 4,09]	0,86	-0,71 [-4,42 ; 3,16]	0,72		
IL8	-2,15 [-5,40 ; 1,21]	0,21	-2,03 [-4,86 ; 0,88]	0,17		
IL10	0,23 [-4,36 ; 5,05]	0,92	-0,60 [-4,58 ; 3,54]	0,77		
IL22	-2,13 [-9,82 ; 6,22]	0,61	-1,06 [-7,84 ; 6,22]	0,77		
TNF	-0,44 [-3,24 ; 2,45]	0,76	-0,53 [-2,96 ; 1,97]	0,68		
IL1	-3,77 [-10,83 ; 3,85]	0,32	-3,85 [-10,18 ; 2,93]	0,26		
IL17a	-0,067 [-5,48 ; 5,66]	0,98	-1,59 [-6,25 ; 3,30]	0,52		

Abréviations : PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; NO₂ = dioxyde d'azote ; O₃ = ozone

Tableau supplémentaire III.B.3.1. Caractéristiques des participants, selon la classe d'IMC (kg/m²) (N = 3074)

	IMC < 25	25 ≤ IMC < 30	IMC ≥ 30
Effectif (%)	1187 (38,61)	1190 (38,71)	697 (22,67)
Age (années), moyenne (ET)	51,43 (7,18)	54,04 (7,13)	54,83 (6,86)
Sexe (hommes), n (%)	412 (34,71)	59,66 (59,66)	355 (50,93)
Non-fumeurs, n (%)	940 (79,19)	973 (81,76)	595 (85,37)
CRPus (mg/L), médiane [IIQ]	0,63 [0,34-1,33]	1,10 [0,63-2,12]	2,10 [1,15-3,99]

Abréviations : IMC = indice de masse corporelle ; CRPus = protéine C-réactive ultrasensible ; ET = écart type ; IIQ = intervalle interquartile

Tableau supplémentaire III.B.3.2. Coefficients de Pearson pour les corrélations entre les taux de polluants atmosphériques résidentiels (µg/m³) en fonction des différentes années à Lille

Corrélations entre les taux de PM ₁₀				
	2010	2011	2012	2013
2010	1,000	0,991	0,997	1,000
2011		1,000	0,997	0,992
2012			1,000	0,998
2013				1,000

Corrélations entre les taux de NO ₂				
	2010	2011	2012	2013
2010	1,000	0,992	0,994	0,999
2011		1,000	0,996	0,991
2012			1,000	0,996
2013				1,000

Abréviations : PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; NO₂ = dioxyde d'azote

Tableau supplémentaire III.B.3.3. Associations entre l'exposition à la pollution atmosphérique à court terme et le taux sérique de CRPus (mg/L) - modèles multipolluants

Population	Effectif	PM ₁₀ ajusté sur NO ₂			PM ₁₀ ajusté sur O ₃		
		Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction
Totale	3011	5,03 [1,13 ; 9,08]	0,011		3,48 [0,52 ; 6,52]	0,021	
Communauté urbaine				0,010			0,073
Lille	1554	4,92 [-0,24 ; 10,34]	0,062		4,52 [0,58 ; 8,62]	0,024	
Dunkerque	1457	3,40 [-2,81 ; 10,00]	0,29		1,37 [-3,25 ; 6,21]	0,57	
Tabac				0,088			0,082
Non-fumeurs	2461	5,03 [0,77 ; 9,47]	0,020		2,17 [-1,01 ; 5,45]	0,18	
Fumeurs actuels	550	6,48 [-2,81 ; 16,66]	0,18		9,72 [2,13 ; 17,87]	0,012	
		NO ₂ ajusté sur PM ₁₀			NO ₂ ajusté sur O ₃		
Population	Effectif	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction
Totale	3011	-2,98 [-7,91 ; 2,22]	0,26		1,24 [-3,22 ; 5,92]	0,59	
Communauté urbaine				0,067			0,032
Lille	1554	1,93 [-6,01 ; 10,54]	0,64		4,60 [-2,66 ; 12,4]	0,22	
Dunkerque	1457	-5,21 [-11,82 ; 1,89]	0,15		-0,54 [-6,43 ; 5,72]	0,86	
Tabac				0,016			0,017
Non-fumeurs	2461	-5,11 [-10,37 ; 0,46]	0,071		-1,78 [-6,48 ; 3,15]	0,47	
Fumeurs actuels	550	5,97 [-6,85 ; 20,56]	0,38		15,37 [2,97 ; 29,26]	0,014	
		O ₃ ajusté sur PM ₁₀			O ₃ ajusté sur NO ₂		
Population	Effectif	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction	Pourcentage de variation pour 10 µg/m ³ [IC 95%]	p	p d'interaction
Totale	3011	0,13 [-2,87 ; 3,23]	0,93		-0,67 [-4,02 ; 2,80]	0,70	
Communauté urbaine				0,039			0,027
Lille	1554	-3,14 [-7,32 ; 1,22]	0,16		-3,36 [-8,19 ; 1,74]	0,19	
Dunkerque	1457	3,98 [-0,54 ; 8,71]	0,085		3,35 [-1,67 ; 8,63]	0,19	
Tabac				0,78			0,83
Non-fumeurs	2461	-0,14 [-3,39 ; 3,22]	0,93		-1,79 [-5,38 ; 1,94]	0,34	
Fumeurs actuels	550	1,04 [-6,39 ; 9,07]	0,79		3,88 [-4,67 ; 13,19]	0,39	

Abréviations : CRPus = protéine C-réactive ultrasensible ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; NO₂ = dioxyde d'azote ; O₃ = ozone

Tableau supplémentaire III.B.3.4. Associations entre l'exposition à la pollution atmosphérique résidentielle et le taux sérique de CRPus (mg/L) - modèles multipolluants

Population	Effectif	PM ₁₀ ajusté sur NO ₂			NO ₂ ajusté sur PM ₁₀		
		Pourcentage de variation pour 2 µg/m ³ [IC 95%]	p	p d'interaction	Pourcentage de variation pour 5 µg/m ³ [IC 95%]	p	p d'interaction
Totale	3074	4,20 [-3,91 ; 12,99]	0,32		-2,39 [-9,61 ; 5,41]	0,54	
Communauté urbaine				0,32			0,13
Lille	1566	5,56 [-6,15 ; 18,73]	0,37		-4,93 [-14,04 ; 5,15]	0,33	
Dunkerque	1508	-4,48 [-25,07 ; 21,75]	0,71		14,39 [-11,35 ; 47,61]	0,30	
Tabac				0,16			0,080
Non-fumeurs	2508	2,47 [-6,23 ; 11,99]	0,59		0,83 [-7,28 ; 9,65]	0,85	
Fumeurs actuels	566	11,70 [-8,33 ; 36,10]	0,27		-16,31 [-30,86 ; 1,31]	0,068	

Abréviations : CRPus = protéine C-réactive ultrasensible ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; NO₂ = dioxyde d'azote

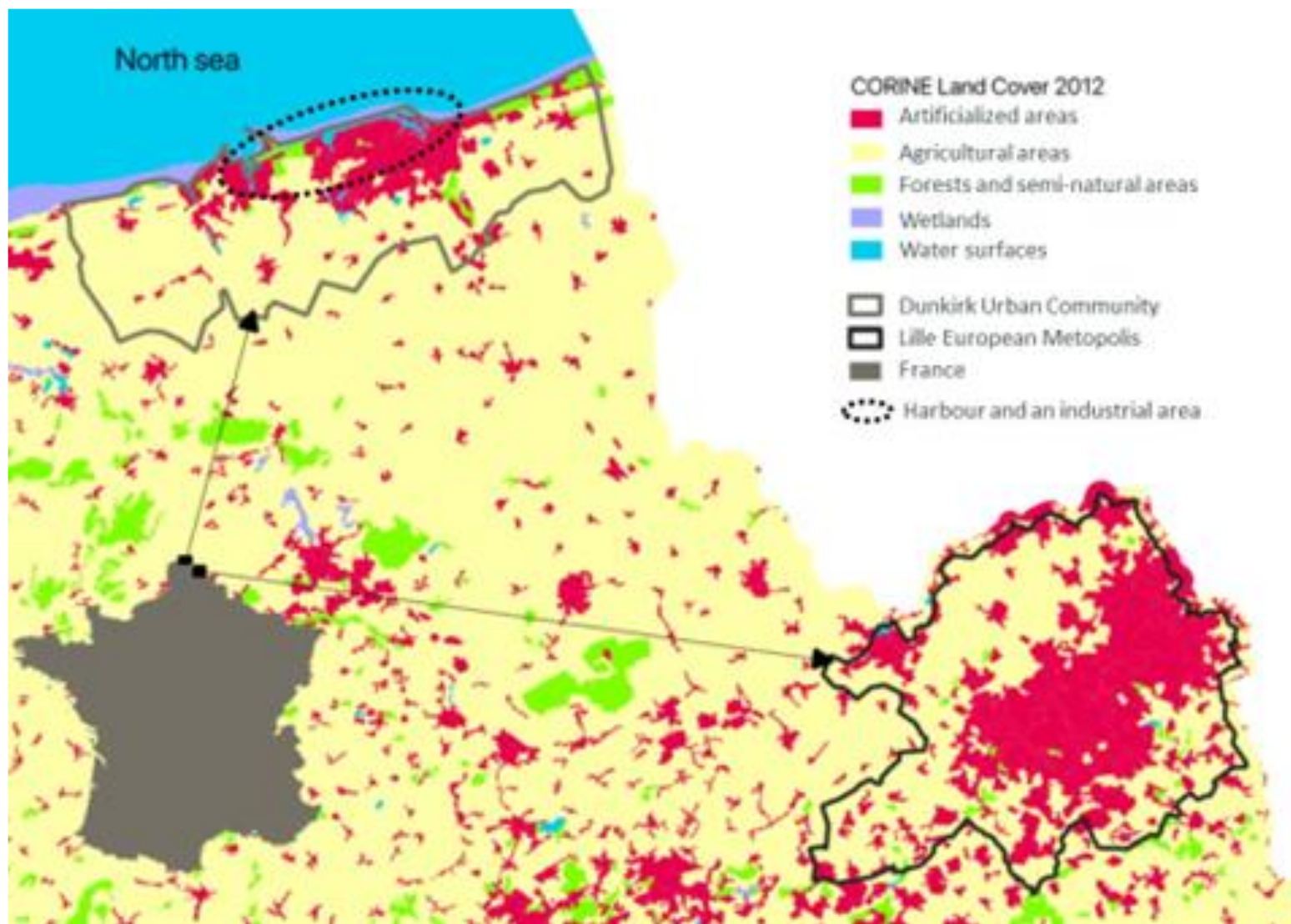


Figure III.B.3.1. Carte du Nord de la France

4. Discussion de cette étude

L'exposition à court terme à la pollution atmosphérique était associée aux concentrations sériques de CRPus, et cette association était influencée par le statut tabagique, l'IMC et la zone urbaine.

L'exposition résidentielle n'était pas significativement associée aux concentrations sériques de CRPus ni aux niveaux plasmatiques de cytokines.

Les associations entre la pollution à court terme et les niveaux de cytokines n'étaient pas significatives, à l'exception de l'association entre l'O₃ et le TNFα.

a) CRPus

Dans une vaste enquête menée auprès d'adultes exposés à des niveaux modérés de pollution atmosphérique dans deux zones urbaines du nord de la France, nous avons constaté qu'une exposition le jour du prélèvement sanguin et la veille était significativement associée à des niveaux plus élevés de CRPus. L'association était plus marquée chez les résidents de Lille, les fumeurs actuels et les personnes non obèses.

Des études antérieures ont mis en évidence des associations dans des populations spécifiques et de plus petite taille (c'est-à-dire comptant moins de 100 personnes). Par exemple, le taux de CRP sanguin était associé à l'exposition professionnelle chez des travailleurs d'une usine de pâte à papier en Suède (138) et chez des agents de police de la circulation au Népal (139). De plus, le taux de CRP dans des taches de sang séché était associé à l'utilisation d'un poêle à bois traditionnel au Nicaragua (156), et le taux sérique de CRP était associé à l'exposition à des fumées métalliques en Allemagne (151). Une association positive significative entre l'exposition quotidienne aux PM_{2,5} et la CRP (mesures répétées) a été observée dans un échantillon de 88 résidents âgés de l'État de l'Utah, aux États-Unis (137).

Les données concernant la population générale sont rares. Deux études menées sur des échantillons de population générale n'ont trouvé aucune association significative. Green *et al.*, (130) ont inclus 2 086 femmes âgées de 42 à 52 ans, issues de six zones urbaines aux États-Unis. Les associations entre les niveaux de polluants atmosphériques (PM_{2,5} et O₃) et la concentration sérique de CRPus n'étaient pas significatives pour des expositions à 1 jour ou 30 jours. De même, une étude portant sur 3 860 membres de la cohorte suisse CoLaus (162) n'a pas mis en évidence d'association significative entre les expositions aux PM₁₀ sur 1 jour, 1 semaine ou 1 mois et la CRPus sérique (avec deux prélèvements sanguins par participant). Enfin, l'étude de Hennig *et al.*, (163) portant sur 8 204 personnes issues de trois grandes villes adjacentes de la région de la Ruhr, en Allemagne, a trouvé une association entre l'exposition à court terme (7 jours) aux PM_{2,5} ou PM₁₀ liée au trafic routier et la CRPus sérique (deux prélèvements par participant). En revanche, les associations avec la pollution d'origine industrielle et avec la pollution totale n'étaient pas significatives.

Il est bien établi que les polluants atmosphériques peuvent provoquer une inflammation pulmonaire directement, par activation des macrophages alvéolaires et sécrétion de cytokines inflammatoires telles que les TNF et les interleukines (IL) (169–171). La surexpression de ces cytokines (principalement l'IL-6) pourrait induire une réponse hépatique aiguë, entraînant une augmentation des concentrations sanguines de diverses protéines associées à l'inflammation systémique, dont la CRP. La concentration sanguine de CRP est habituellement faible dans des conditions normales, mais elle augmente rapidement et de manière marquée en réponse à un stimulus pro-inflammatoire (172–174). La demi-vie plasmatique de la CRP est inférieure à 20

h et ne varie pas selon les conditions physiopathologiques. Ainsi, le seul déterminant de la concentration circulante de CRP est son taux de synthèse, qui reflète directement l'intensité du stimulus. Les résultats de plusieurs études *in vitro* et *in vivo* suggèrent que les niveaux de CRP augmentent en réponse à une exposition aux PM (175–182).

Les effets de la pollution atmosphérique pourraient varier selon la durée d'exposition. En effet, une exposition à court terme pourrait déclencher la phagocytose des polluants atmosphériques par les macrophages alvéolaires, ce qui pourrait entraîner une augmentation soudaine des niveaux de biomarqueurs inflammatoires (169). Une exposition à long terme pourrait, quant à elle, entretenir le processus inflammatoire, avec des conséquences sur les maladies chroniques (162).

Plusieurs des interactions étudiées avec les polluants étaient statistiquement significatives. Tout d'abord, nous avons observé une interaction significative entre le statut tabagique et l'exposition à court terme au NO₂. Cette interaction persistait dans les modèles multipolluants. Nos résultats suggèrent que l'effet de la pollution sur les niveaux de CRPus est plus important chez les fumeurs que chez les non-fumeurs.

Pilz *et al.*, (152) ont également observé une interaction significative entre la pollution et le tabagisme. Toutefois, il est difficile d'en tirer des conclusions fermes, car l'association était plus forte chez les anciens fumeurs que chez les non-fumeurs et les fumeurs actuels. La plus forte association observée chez les fumeurs pourrait être liée à des effets synergiques entre la pollution atmosphérique et le tabac. Une hypothèse est que les fumeurs sont plus sensibles aux effets de la pollution car leurs voies respiratoires et leurs poumons sont déjà endommagés. En effet, le tabagisme entraîne une altération des jonctions serrées et affaiblit ainsi la fonction de barrière épithéliale des voies respiratoires (183). De plus, une perméabilité épithéliale accrue chez les fumeurs pourrait faciliter la pénétration des polluants dans l'organisme (184,185). Par ailleurs, une diminution de la clairance pulmonaire chez les fumeurs pourrait expliquer l'effet plus marqué des polluants (186). En effet, une étude a montré que les personnes exposées à la fumée de cigarette ou fumant activement (187,188) présentaient des niveaux plasmatiques d'antioxydants significativement plus faibles, ce qui pourrait rendre ces personnes plus sensibles au stress oxydatif généré par la pollution atmosphérique (189). De manière cohérente, Sun *et al.*, (190) ont rapporté que le risque de mortalité lié à la pollution atmosphérique était plus élevé chez les fumeurs actuels ou anciens que chez les non-fumeurs. Cependant, un effet synergique n'est pas systématiquement observé. Panasevich *et al.*, (191) ont trouvé une association significative entre la pollution atmosphérique et la CRPus chez les non-fumeurs, mais pas chez les fumeurs. Une hypothèse est que les fumeurs présentent déjà un niveau élevé d'inflammation de bas grade en raison des particules issues du tabac, ce qui pourrait rendre plus difficile la détection d'un effet additionnel de la pollution atmosphérique chez les fumeurs par rapport aux non-fumeurs. Ainsi, les données disponibles sur l'interaction entre tabac et pollution atmosphérique sont limitées et non concluantes. Cette interaction mérite donc d'être étudiée plus en profondeur.

Deuxièmement, nous avons observé une interaction significative entre les polluants (PM₁₀ et NO₂) et l'IMC. Ainsi, les PM₁₀ étaient positivement associées à la CRPus chez les personnes non obèses mais pas chez les personnes obèses. Ces résultats sont cohérents avec ceux de Pilz *et al.*, (152), qui rapportaient une association entre des niveaux élevés de CRPus et les PM₁₀ chez les individus non obèses mais pas chez les obèses. Dans notre étude (et comme attendu au vu des résultats de Gentile (192), le niveau de CRPus était beaucoup plus élevé chez les personnes obèses que chez celles de poids normal. Cela est probablement lié au fait que l'obésité est associée à une inflammation chronique, à l'activation de voies pro-inflammatoires et à des

niveaux plus élevés de biomarqueurs inflammatoires (193). Ainsi, un effet modéré de la pollution atmosphérique pourrait être plus difficile à détecter dans une population présentant déjà un niveau de base très élevé d'inflammation de bas grade.

À l'inverse, il a été suggéré que l'effet de la pollution atmosphérique sur la santé (en dehors de l'inflammation de bas grade) pourrait être plus important chez les personnes obèses. Par exemple, dans la cohorte Danish Nurse (194), l'association entre les $PM_{2,5}$ et le diabète était plus forte chez les personnes obèses. De manière surprenante, nous avons trouvé que l'exposition au NO_2 était associée à des niveaux significativement plus faibles de CRPus chez les personnes obèses. Nous n'avons pas d'hypothèse physiologique pour expliquer ce résultat, qui pourrait être lié aux corrélations négatives entre NO_2 et O_3 et aux effets délétères de l' O_3 . Néanmoins, l'association entre O_3 et CRPus n'était pas significative et ne permet pas de confirmer cette hypothèse. Ce résultat pourrait également être dû au hasard.

Troisièmement, nous avons observé des interactions entre l'exposition à court terme au NO_2 et à l' O_3 et la zone urbaine. L'association entre l'exposition au NO_2 et à l' O_3 et la CRPus était statistiquement significative à Lille mais pas à Dunkerque, avec une interaction significative. L'association entre l' O_3 et la CRPus était de signe opposé à celle entre le NO_2 et la CRPus. En effet, les niveaux d' O_3 et de NO_2 sont négativement corrélés.

Il existe une zone industrielle au nord-ouest de l'agglomération de Dunkerque ; ainsi, l'interaction entre les niveaux d' O_3 et de NO_2 et la zone urbaine pourrait s'expliquer par une différence dans la nature des polluants entre Lille et Dunkerque. Dans ce contexte, l'association observée dans l'étude de Hennig *et al.*, (163) — où les pollutions d'origine industrielle et routière étaient modélisées séparément — n'était significative que pour la pollution liée au trafic routier. Un vent de secteur nord aurait tendance à amener la pollution industrielle vers Dunkerque, tandis qu'un vent de secteur sud amènerait la pollution liée au trafic routier vers la ville, reproduisant ainsi une situation similaire à celle de Lille.

Avec un vent de secteur sud, les associations entre pollution et inflammation devraient donc être similaires à Lille et à Dunkerque. Avec un vent de secteur nord, la nature des polluants serait plus différente. On pourrait donc raisonnablement émettre l'hypothèse d'une interaction pollution/vent à Dunkerque. Cependant, nous n'avons pas observé d'interaction significative en ce sens. La nature des polluants à Lille et à Dunkerque et les différences entre les deux zones restent donc à explorer davantage. Enfin, l'interaction entre le NO_2 et la zone urbaine n'était plus significative après prise en compte de l'interaction avec l'IMC, ce qui suggère que l'association plus faible observée à Dunkerque (par rapport à Lille) était en partie liée à un IMC moyen plus élevé dans la première ville.

Nous n'avons pas trouvé d'association significative entre l'exposition résidentielle et le niveau de CRPus. Cela pourrait s'expliquer par le fait que la plupart des effets de la pollution atmosphérique sont de courte durée. Autre explication possible : l'écart entre le niveau le plus élevé et le plus faible d'exposition aux PM_{10} dans notre population modérément exposée pourrait être trop faible pour produire une différence significative. Ainsi, l'effet de la pollution atmosphérique sur l'inflammation (s'il existe) pourrait ne pas varier fortement d'une zone à l'autre au sein d'une même ville. Enfin, les politiques environnementales tendent à viser la réduction globale des niveaux de pollution et non uniquement la diminution du nombre de sujets très exposés.

Dans la littérature, des associations positives entre exposition professionnelle à la pollution et CRP ont été trouvées chez 87 conducteurs de véhicules automobiles au Pakistan (140), 39 chauffeurs de taxi au Brésil (143), 72 employés dans une station-service en Irak (142), 137 personnes exposées aux fumées de moteurs diesel en Chine (144) et 78 soudeurs en Suède

(150). À l'inverse, une association négative significative entre exposition professionnelle et CRP a été observée dans une étude chinoise portant sur 54 travailleurs exposés aux fumées de moteurs diesel, par rapport à 55 participants non exposés (145).

Certaines études ont trouvé une association entre pollution atmosphérique et niveau de CRP dans des populations particulières, comme 175 patients dialysés à Taïwan (141), 6 508 femmes enceintes aux Pays-Bas (157) et 7 915 personnes âgées en Chine (160). Par ailleurs, une association a été mise en évidence entre la densité du trafic routier et le niveau de CRP chez 1 017 participants d'origine portoricaine vivant à Boston, États-Unis (131).

Plusieurs études menées en population générale — par exemple, 3 860 participants en Suisse (162), 408 participants à Boston (États-Unis) (161), 2 252 participants en Allemagne (152) et environ 17 000 participants au Royaume-Uni (153) — n'ont pas trouvé d'association significative entre l'exposition à long terme à la pollution atmosphérique et la CRP. Cependant, l'étude de Green *et al.*, (130) portant sur 2 086 femmes aux États-Unis a trouvé une association significative entre la CRPus et l'exposition annuelle aux $PM_{2,5}$, ainsi qu'avec l'exposition sur 6 mois à l' O_3 . Enfin, une association significative entre la CRPus et l'exposition annuelle aux $PM_{2,5}$ ou aux PM_{10} a été trouvée chez 8 204 participants allemands (163). Dans cette étude, la concentration résidentielle moyenne sur 365 jours de PM_{10} était de $19,68 \mu\text{g}/\text{m}^3$ pour la pollution totale, dont une moyenne de $0,81 \mu\text{g}/\text{m}^3$ provenant de la pollution routière et $2,47 \mu\text{g}/\text{m}^3$ de la pollution industrielle. Il est à noter que ces deux études comportaient des mesures répétées du niveau de CRPus.

b) Cytokines

Les associations entre les niveaux plasmatiques de cytokines et les expositions à court terme ou résidentielles n'étaient pas significatives, à l'exception de l'association entre l'exposition à court terme à l' O_3 et le $TNF\alpha$.

Des études antérieures ont montré des associations entre la pollution atmosphérique et les cytokines chez des personnes fortement exposées ou dans des populations spécifiques. Par exemple, des études menées auprès de personnes professionnellement exposées à la pollution ont trouvé des associations positives (140,142,143,145) ou négatives (144,154,155) entre l'exposition et les niveaux de cytokines pro-inflammatoires.

Des études menées chez des enfants asthmatiques (146,149,159) ont mis en évidence des associations positives entre l'exposition à la pollution et les niveaux sanguins de cytokines pro-inflammatoires et, de manière surprenante dans deux publications, de la cytokine anti-inflammatoire IL-10 (146,159). L'exposition aux PM_{10} était positivement corrélée à l'IL-1 β mais négativement corrélée à l'IL-6 dans une cohorte de nouveau-nés (158) et positivement corrélée à l'IL-6 mais négativement au $TNF\alpha$ chez des femmes enceintes (147).

Dans des études interventionnelles, l'exposition à des fumées métalliques était associée à des niveaux plus élevés de cytokines (151), et l'utilisation d'un purificateur d'air était associée à des niveaux plus faibles d'IL-6 (148).

À l'inverse, les études en population générale ont publié peu de données concernant des niveaux modérés d'exposition. Lane *et al.*, (161) ont étudié 408 résidents de Boston mais n'ont pas trouvé d'association significative entre l'exposition annuelle moyenne à la pollution atmosphérique et les niveaux de cytokines. Enfin, Tsai *et al.*, (162) ont étudié une cohorte de 3 860 sujets âgés de 35 à 75 ans en Suisse. Les mesures des biomarqueurs ont été répétées deux fois, et l'exposition aux PM_{10} a été estimée à partir de l'adresse du domicile des participants. L'exposition à court terme a été définie de différentes façons : exposition un jour,

une semaine ou un mois avant le prélèvement sanguin. En moyenne, les concentrations de PM₁₀ étaient de 22,2 µg/m³ un jour avant le prélèvement sanguin, 24,5 µg/m³ une semaine avant et 24,8 µg/m³ un mois avant.

L'exposition à long terme était définie comme l'exposition survenue 3 mois et 6 mois avant le prélèvement sanguin. Les concentrations de polluants observées dans l'étude de Tsai *et al.*, étaient similaires à celles de notre étude. Tsai *et al.*, ont trouvé une association statistiquement significative entre l'exposition à court et à long terme aux PM₁₀ et les niveaux de trois cytokines (IL-1β, IL-6 et TNFα) ; il s'agit à ce jour de la seule étude ayant mis en évidence une association avec les niveaux de cytokines dans la population générale.

Il est plausible que la pollution atmosphérique augmente l'inflammation et la sécrétion de cytokines, car des études *in vitro* ont montré la production de TNFα par des macrophages alvéolaires exposés à des particules atmosphériques (169,170). De plus, l'exposition à la pollution de l'air peut entraîner une augmentation du stress oxydatif (195), une incapacité à éliminer les cellules apoptotiques et un état d'inflammation chronique (196,197).

L'absence de résultats significatifs concernant l'IL-1β et l'IL-6 semble discordante avec les résultats de Tsai *et al.*, Concernant la taille de l'effet observée par Tsai *et al.*, (162), la puissance statistique de notre étude pour l'IL-1β était de 99 %. Par ailleurs, lorsque l'on compare les associations pour l'IL-6 dans l'étude de Tsai *et al.*, et dans la nôtre, les intervalles de confiance ne se chevauchent pas — ce qui suggère une réelle différence entre les deux séries de résultats. En effet, la valeur observée était d'environ 0,049 [0,029 ; 0,067] unité de variation par incrément de 1 µg/m³ de PM₁₀ dans la cohorte suisse, contre -0,0036 [-0,025 ; 0,018] dans notre étude.

Contrairement à Tsai *et al.*, nous avons observé une association positive entre l'O₃ et le TNFα. Tsai *et al.*, avaient trouvé une association positive entre le TNFα et les PM₁₀. Dans la plupart des études, les niveaux de PM₁₀ et de NO₂ sont négativement corrélés aux niveaux d'O₃ ; ainsi, nous nous attendions à observer une association positive entre le TNFα et les PM₁₀ et NO₂, et une corrélation négative avec l'O₃. Toutefois, ce résultat doit être interprété avec prudence, car nous avons effectué de nombreux tests statistiques sur les données relatives aux cytokines. Enfin, nos résultats majoritairement non significatifs concernant les cytokines ne sont pas cohérents avec ceux de Tsai *et al.*, malgré un niveau de puissance statistique apparemment suffisant. Néanmoins, le fait que nous ayons analysé un seul prélèvement sanguin pourrait également expliquer la divergence entre les deux séries de résultats. Ainsi, d'autres études sont nécessaires pour explorer l'association présumée entre pollution atmosphérique et niveaux de cytokines.

Forces et limites

La présente étude présente plusieurs points forts.

Tout d'abord, nous avons étudié un large échantillon en population générale. Ensuite, nos dosages de biomarqueurs ont été réalisés avec des procédures strictes de contrôle qualité. En particulier, une analyse à haut débit des niveaux plasmatiques de cytokines a été réalisée sur une plateforme de dosage multiplex équipée d'un système de détection par électrochimiluminescence haute performance (QuickPlex SQ 120, Meso Scale Diagnostics LLC). Comme décrit précédemment (198,199), cet instrument QuickPlex est supérieur aux systèmes ELISA multiplex classiques en termes de sensibilité, de reproductibilité, de débit et de capacité à tester simultanément des antigènes étroitement apparentés sans réduction de signal liée à la compétition antigénique.

Notre étude comporte toutefois certaines limites. Premièrement, nous ne disposons pas de données d'exposition résidentielle pour 2011. Cependant, le coefficient de corrélation de Pearson entre les niveaux d'exposition résidentielle en 2012 et en 2013 à Dunkerque était de 0,94 pour les PM₁₀ et de 0,97 pour le NO₂ ; et la corrélation intraclasse pour les données imputées de 2012 à partir des données de pollution de 2013 était de 0,94 pour les PM₁₀ et de 0,89 pour le NO₂ — ce qui suggère que la distribution géographique des niveaux annuels moyens de pollution atmosphérique est stable dans le temps et que l'imputation des données d'exposition résidentielle à Dunkerque pourrait n'induire qu'un biais de classification limité. De plus, l'ajustement de l'analyse sur l'année d'inclusion pourrait avoir réduit ce biais. Par ailleurs, telle que définie dans notre étude, l'exposition résidentielle dépendait principalement de la localisation géographique du foyer et moins du moment de la mesure. Ainsi, le biais lié à l'utilisation de données sur les disparités géographiques de 2012 (plutôt que 2011) pourrait être limité.

Deuxièmement, la différence d'exposition résidentielle entre les participants les plus et les moins exposés était faible. Par exemple, l'intervalle interquartile était de 25,84–27,19 µg/m³ à Lille, ce qui a pu limiter la puissance de l'étude pour l'analyse de l'exposition résidentielle.

Troisièmement, un manque de précision dans la mesure des concentrations de polluants pourrait avoir induit un biais non différentiel et réduit la puissance de l'étude.

Quatrièmement, le caractère transversal de notre étude (avec un seul prélèvement sanguin) ne nous a pas permis de contrôler la variabilité intra-individuelle des niveaux de cytokines. En comparaison, des études avec mesures répétées des biomarqueurs ont montré des associations significatives entre l'exposition à la pollution et les cytokines (162) ou la CRPus (163).

Néanmoins, nos résultats statistiquement significatifs concordaient avec des associations déjà connues (200–202), comme les associations avec l'IMC, l'IL-6 ($p < 10^{-17}$), le TNF ($p < 10^{-3}$) et l'association entre l'obstruction pulmonaire et l'IL-6 ($p < 0,05$) (données non présentées). Cela confirme notre capacité à mettre en évidence une forte association, même lorsque les niveaux de cytokines présentent une variabilité importante.

Enfin, la taille de notre échantillon pourrait ne pas avoir été suffisante pour certaines analyses (en particulier l'analyse des cytokines). Par exemple, l'association entre la pollution et la CRPus dans le sous-échantillon de 982 non-fumeurs de Lille n'était pas significative (données non présentées), ce qui suggère la nécessité d'un échantillon plus large.

c) Conclusion de cette étude

Nous avons observé une association significative entre l'exposition à court terme à la pollution atmosphérique et le niveau sérique de CRPus ; cette association est cohérente avec l'hypothèse selon laquelle les effets sanitaires de la pollution atmosphérique sont médiés par une inflammation de bas grade. Ce résultat concorde avec certaines observations antérieures (mais pas toutes).

Nous avons constaté que l'IMC et le tabagisme interagissaient avec l'association entre l'exposition à court terme à la pollution atmosphérique et le niveau sérique de CRPus. Cette interaction pourrait expliquer certaines divergences dans les données de la littérature et devrait être prise en compte dans les études portant sur la CRPus et la pollution.

Nous avons étudié un large panel de cytokines ; toutes les associations entre l'exposition à la pollution atmosphérique et les niveaux de cytokines étaient non significatives, à l'exception d'une seule, et ne concordaient donc pas avec les résultats antérieurement rapportés par Tsai *et al.*, (162). Ainsi, des recherches supplémentaires sur ce sujet sont nécessaires, incluant des

mesures précises des polluants atmosphériques et tenant compte de la variabilité intra-individuelle des niveaux de cytokines.

C. Revue de la littérature :

1. Introduction

La métabolomique permet d'identifier des liens entre des expositions environnementales et des perturbations physiologiques infracliniques. Ainsi, elle permet de mettre en évidence de faibles effets, associés à une exposition modérée. Ces faibles effets peuvent être sans manifestations cliniques visibles mais pourraient avoir un impact à long terme sur la santé.

Par ailleurs, la métabolomique permet la mesure de centaines voire de milliers de molécules (203). Ce grand nombre de variables entraîne une problématique spécifique pour les analyses en raison de la multiplicité des tests statistiques.

De plus, les récents progrès technologiques et la diminution des coûts ont permis de réaliser des analyses métabolomiques sur un plus grand nombre d'échantillons, ce qui élargit la gamme des analyses statistiques envisageables. Les analyses multidimensionnelles ne sont plus les seules analyses utilisées. En effet, des analyses portant sur des métabolites individuels avec des corrections pour tests multiples sont désormais réalisables.

Néanmoins, il est nécessaire d'avoir un consensus sur les méthodes utilisables, afin de pouvoir comparer les études entre elles.

À ce jour, il n'existe pas de recommandations bien établies concernant les méthodes statistiques à utiliser pour étudier les associations entre exposition environnementale et métabolomique dans les études par chromatographie couplée à la spectrométrie de masse (CSM).

Il est donc nécessaire de mener une réflexion quant aux différentes stratégies statistiques employées afin d'établir des recommandations plus claires.

L'objectif de cette revue était d'examiner systématiquement les objectifs spécifiques et les méthodes statistiques utilisées dans les études de grande taille évaluant les associations entre les expositions environnementales et les modifications métabolomiques (issues de la chromatographie-spectrométrie de masse, liquide ou gazeuse). Nous avons cherché à identifier les questions scientifiques abordées, à évaluer la pertinence des hypothèses statistiquement testables correspondantes (H1), et à apprécier la validité des analyses statistiques.

Les expositions environnementales à inclure étaient définies comme relevant des environnements comportementaux, sociaux, naturels et physiques.

Dans ce travail, nous ne nous concentrons pas sur les questions techniques liées aux analyses biochimiques (72,204).

2. Méthode

Nous avons effectué une recherche bibliographique dans les bases de données PubMed, Embase et Web of Science, afin d'identifier les publications en langue anglaise disponibles jusqu'en avril 2024. Les études non publiées n'ont pas été prises en compte. Les équations de recherche étaient les suivantes:

Pour PubMed : (("Metabolomic" OR "Metabolomics" OR "Metabolomics"[Mesh]) AND ("Case-control" OR "Cross-sectional" OR "Epidemiological" OR "Cohort" OR "Case-Control Studies"[Mesh] OR "Cross-Sectional Studies"[Mesh] OR "Epidemiologic Studies"[Mesh] OR "Cohort Studies"[Mesh]) AND ("Liquid chromatography" OR "Gas chromatography" OR "Mass spectrometry" OR "High-resolution mass spectrometry" OR "LC-HRMS" OR "UPLC-MS" OR "GC-MS" OR "HRMS" OR "Chromatography, Liquid"[Mesh] OR "Mass Spectrometry"[Mesh] OR "Liquid Chromatography-Mass Spectrometry"[Mesh] OR "Chromatography, Gas"[Mesh] OR "Gas Chromatography-Mass Spectrometry"[Mesh])).

Pour Web of Science : (("Metabolomic" OR "Metabolomics") AND ("Case-control" OR "Cross-sectional" OR "Epidemiological" OR "Cohort" OR "Case-Control Studies" OR "Cross-Sectional Studies" OR "Epidemiologic Studies" OR "Cohort Studies") AND ("Liquid chromatography" OR "gas chromatography" OR "Mass spectrometry" OR "High-resolution mass spectrometry" OR "LC-HRMS" OR "UPLC-MS" OR "GC-MS" OR "HRMS" OR "Chromatography, Liquid" OR "Liquid Chromatography-Mass Spectrometry" OR "Chromatography, Gas" OR "Gas Chromatography-Mass Spectrometry")).

Pour Embase : (('metabolomic' OR 'metabolomics' OR 'metabolomics'/exp) AND ('case-control' OR 'cross-sectional' OR 'epidemiological' OR 'cohort' OR 'case control study'/exp OR 'cross-sectional study'/exp OR 'epidemiology'/exp OR 'cohort analysis'/exp) AND ('liquid chromatography' OR 'gas chromatography' OR 'mass spectrometry' OR 'high-resolution mass spectrometry' OR 'lc-hrms' OR 'uplc-ms' OR 'gc-ms' OR 'hrms' OR 'liquid chromatography'/exp OR 'mass spectrometry'/exp OR 'liquid chromatographymass spectrometry'/exp OR 'gas chromatography'/exp OR 'mass fragmentography'/exp)) AND 'human'/de NOT ('animal cell'/de OR 'animal experiment'/de OR 'animal model'/de OR 'animal tissue'/de OR 'case report'/de OR 'cell culture'/de OR 'controlled study'/de OR 'diagnostic test accuracy study'/de OR 'human cell'/de OR 'in vitro study'/de OR 'in vivo study'/de OR 'pilot study'/de OR 'quality control'/de OR 'randomized controlled trial'/de OR 'validation process'/de OR 'validation study'/de) AND ('article'/it OR 'conference abstract'/it).

En complément, nous avons appliqué une approche dite « boule de neige » et utilisé l'outil Biblizap pour examiner les listes de références des publications incluses (205).

Les critères d'inclusion étaient les suivants : un article original, l'utilisation de la chromatographie couplée à la spectrométrie de masse (CSM), la présence de données métabolomiques, une taille d'échantillon d'au moins 500 individus (seuil à partir duquel (i) un éventail plus large d'analyses peut être appliqué et (ii) la correction pour tests multiples devient sujette à débat), une approche métabolomique non ciblée (c'est-à-dire des approches mesurant les métabolites sans hypothèses a priori, sans se focaliser sur des métabolites prédéfinis), et l'évaluation des associations entre une exposition environnementale (par exemple pollution, exposition professionnelle, alimentation ou tabac) et la métabolomique.

Les critères d'exclusion étaient les suivants : une exposition liée à une maladie, aux médicaments ou au microbiome, les études *in vitro*, les études animales, les articles de revue, les études interventionnelles, une approche métabolomique ciblée, et les études dont l'unique objectif était l'identification de proxys d'exposition. En effet, nous nous sommes concentrés sur les études utilisant la métabolomique pour comprendre les liens entre l'environnement et le phénotype, c'est-à-dire la relation entre exposition et métabolomique dans une perspective

de caractérisation des effets biologiques de l'environnement. Nous avons donc exclu les études dont l'objectif unique était l'identification de proxys d'exposition puisque leur objectif était de mesurer indirectement l'exposition plutôt que de fournir des explications physiopathologiques.

Deux lecteurs (Marion Darras-Hostens et Luc Dauchet) ont extrait les données et procédé indépendamment au tri des titres et résumés. Les articles retenus à cette étape ont ensuite été lus dans leur intégralité. Le processus de sélection a été documenté à l'aide du logiciel Rayyan (206). L'étude a été enregistrée dans la base de données PROSPERO le 28 mai 2024. Les principaux critères de jugement de cette revue étaient les objectifs spécifiques et les méthodes d'analyse statistique utilisées pour étudier les associations entre les expositions environnementales et la métabolomique fondée sur la CSM. Les variables suivantes ont été extraites de chaque publication incluse : premier auteur, année de publication, pays, période de l'étude, type d'étude (cohorte, cas-témoins ou transversale), objectifs (en lien avec la méthode utilisée pour y répondre), population (taille d'échantillon, méthode de recrutement, âge (moyenne \pm écart-type (ET), médiane [intervalle interquartile (IIQ)]), sexe (% hommes/femmes/autres), et caractéristiques (population générale ou population spécifique)), exposition (types d'exposition, quantification sous forme de moyenne \pm ET, médiane [IIQ], ainsi que la durée), variables métabolomiques (type de fluide biologique, nombre de métabolites mesurés, nombre de métabolites annotés), méthode(s) statistique(s) utilisée(s), et représentation(s) graphique(s) des résultats. Nous avons utilisé les cinq premiers items du score ROBINS-I pour évaluer le risque de biais (207).

De plus, nous avons déterminé la question scientifique de chaque étude. Cette question correspond au gain de connaissance attendu, par exemple en termes d'amélioration de la compréhension physiopathologique ou d'impact en santé publique. La question scientifique a été déterminée uniquement à partir de l'introduction, puis regroupée en catégories. Une classification a posteriori a été établie après inclusion et lecture complète des études.

Nous avons recensé les hypothèses testées (H1) uniquement à partir de la dernière phrase de l'introduction de chaque étude, qui correspond aux objectifs explicitement énoncés de l'étude. Les hypothèses H1 et H0 pouvant formellement correspondre à des tests statistiques sont rarement explicitement formulées dans les articles d'épidémiologie. Le meilleur proxy identifié pour ces hypothèses est constitué par les objectifs principaux et secondaires.

Nous n'avons pas utilisé la section méthodologie pour définir l'objectif général et les hypothèses H1, afin d'éviter de reconstruire une hypothèse scientifique à partir de la méthodologie statistique, et de conclure à tort à la cohérence de celle-ci.

Nous avons déterminé les outcomes statistiques (par exemple, quelles associations étaient testées entre une variable à expliquer et une variable explicative) et identifié les tests statistiques à partir de la section statistique de la méthodologie.

Enfin, nous avons évalué (i) si chaque hypothèse H1 était appropriée à la question scientifique correspondante. Sur la base des analyses statistiques réalisées, nous avons ensuite déterminé (ii) si l'outcome statistique était approprié à son hypothèse H1 correspondante, et (iii) la validité des tests statistiques ; autrement dit, nous avons évalué si un problème statistique identifié pouvait avoir affaibli la validité des p-values rapportées.

Enfin, nous avons considéré que le terme « profil », employé dans certaines publications, désigne une description du métabolome prenant en compte des variables multiples (par exemple, les dimensions dérivées d'une ACP ou d'une PLS, ou encore les classes issues d'une classification hiérarchique), plutôt que des métabolites individuels ou des voies métaboliques.

3. Résultats de cette étude

a) Sélection des publications

Les recherches dans les bases de données ont permis d'identifier un total de 5858 références (figure III.C.3.1). Après exclusion des doublons et sélection des articles sur la base du titre et du résumé puis du texte intégral, 18 articles ont été retenus à ce stade. L'application d'une stratégie de type « boule de neige » aux listes de références de ces 18 études a permis d'inclure 6 publications supplémentaires. Au total, 24 publications (portant sur 22 populations ou échantillons de population distincts) ont donc été incluses à la revue (tableau III.C.3.1 et tableau supplémentaire III.C.3.1).

Les articles exclus à l'étape de lecture intégrale concernaient, par exemple, des études ayant recours à une approche métabolomique ciblée (208,209), ou des études dont l'unique objectif était l'identification de proxies d'exposition (210,211).

La taille des échantillons variait de 500 à 16 222 individus. Concernant le score de risque de biais ROBINS-I calculé pour chaque étude, 11 études présentaient un « faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels » (64,212–221) et 13 présentaient « certaines préoccupations » (65,66,222–232) (tableau III.C.3.2).

b) Questions scientifiques

Selon les articles inclus, nous avons identifié deux principales questions scientifiques (Tableau III.C.3.3 et Tableau supplémentaire III.C.3.2) :

1) « Quels processus ou voies métaboliques spécifiques sont influencés par l'exposition ? »
Les articles dont le contexte contenait des phrases telles que « *may provide insight into the biological mechanisms* » ou « *gain more in-depth knowledge about the biochemical pathways* » ont été inclus dans cette catégorie.

2) « Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ? » En statistique, un modèle de médiation vise à identifier et à expliquer le mécanisme sous-jacent à la relation entre une variable indépendante et une variable dépendante en introduisant une troisième variable, appelée médiateur (également appelée variable intermédiaire). Ainsi, les études qui cherchaient à examiner les associations entre une exposition et une variable métabolomique, puis entre une variable métabolomique et un phénotype, ont été incluses dans la catégorie « médiation ».

Treize des 24 études avaient un objectif classé comme « Quels processus ou voies métaboliques spécifiques sont influencés par l'exposition ? »

(64,212,213,215,217,218,222,225–228,230,231). De plus, deux de ces études avaient un objectif supplémentaire consistant à tester des interactions (215,218).

Par ailleurs, 10 études (65,66,216,219–224,229) visaient à étudier les associations entre une exposition et une variable métabolomique, puis entre une variable métabolomique et un

phénotype ; logiquement, ce type d'objectif devrait correspondre à une approche de médiation. Nous avons donc classé ces articles dans la catégorie « Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ? »

En outre, deux études utilisaient les termes « profil » (214) et « signature » (65) sans les définir clairement, et l'une de leurs questions scientifiques n'a pas pu être classée.

Enfin, une étude présentait un objectif qui n'a pas pu être classé sur la base des éléments de l'introduction (232).

Il convient de noter que certaines études cherchaient également à identifier des biomarqueurs ou des proxies d'exposition, en plus de la question scientifique répondant à nos critères d'inclusion (213,214,224).

c) Hypothèses scientifiques

Les hypothèses H1 définies sont présentées dans le Tableau III.C.3.3 et dans le Tableau supplémentaire II.C.3.2. Dix études avaient une hypothèse H1 classée comme suit : « Il existe une association entre au moins un métabolite et une exposition environnementale » (64,212,213,216–218,222,225,230,232), et 4 études avaient l'hypothèse H1 : « Il existe une association entre au moins une voie métabolique et une exposition environnementale » (225,227,231,232).

Lorsque l'objectif énoncé dans l'étude incluait à la fois une association entre l'exposition et les métabolites, puis entre les métabolites et le phénotype, nous avons considéré que l'hypothèse H1 concernait la médiation. Par ailleurs, 7 études avaient l'hypothèse H1 : « Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype » (66,219,221–224,229), 1 étude avait l'hypothèse H1 : « Au moins une voie métabolique est un médiateur dans la relation entre l'exposition et le phénotype » (66) et 1 étude avait l'hypothèse H1 : « Au moins un score métabolique est un médiateur dans la relation entre l'exposition et le phénotype » (223). Enfin, deux études avaient pour hypothèse H1 : « Il existe une interaction avec l'IMC » (216,218).

Certaines études utilisaient des termes tels que « profil » (214,215,226,228), « signature » (65), ou « métabolome » (220) dans l'énoncé de leur objectif, sans toutefois en fournir une définition claire, ce qui rendait impossible la formulation précise de l'hypothèse H1. En effet, ces termes peuvent renvoyer à plusieurs types de variables statistiques, tels que des métabolites individuels (65,214,215,220,226,228), des voies métaboliques (65,215), un score utilisé pour une analyse de médiation (215,226), ou encore des clusters (220). Il ne semble pas exister de consensus sur la signification précise de ces termes ; par conséquent, nous n'avons pas pu classer l'hypothèse H1 ni l'objectif sur la seule base de ces termes, sans clarification supplémentaire.

Enfin, il convient de noter que les études dont la question scientifique portait sur l'identification de biomarqueurs ou de proxies (213,214,224) n'étaient pas concernées par la formulation d'une hypothèse H1.

d) Exposition et critère de jugement

Fluides

Le type d'échantillon analysé était le sang dans 23 études (64–66,212–219,221–232), le sang et le liquide céphalorachidien dans une étude (219), et les cheveux dans une autre étude (220) (tableau supplémentaire III.C.3.3).

Exposition

Différentes expositions environnementales ont été étudiées: 12 études ont porté sur les habitudes alimentaires ou de consommation de boissons (212–214,216,218,222,223,225–227,229,230), 7 sur des expositions liées aux modes de vie (par exemple, le tabagisme ou le lieu d'habitation) (65,217,219,221,224,231,232), 4 sur des polluants (64,66,220,228), et une sur les abus précoces c'est-à-dire la maltraitance physique et sexuelle avant l'âge de 18 ans (215) (tableau III.C.3.1 et tableau supplémentaire III.C.3.3).

e) Analyses statistiques

(1) Méthodes d'analyses unidimensionnelles

Analyses de l'ensemble des métabolites comparées à celles limitées aux métabolites annotés

Quatorze articles ont analysé uniquement les métabolites annotés (64,65,212,214,216–218,220,222–224,226–228), tandis que les 10 autres ont analysé l'ensemble des métabolites mesurés (66,213,215,219,221,225,229–232).

Analyses sur les métabolites individuels

Dix-neuf études ont évalué les associations entre les expositions environnementales et des métabolites individuels (64–66,215–221,223–228,230–232) (tableau III.C.3.1). Toutes ont appliqué une correction pour tests multiples, en utilisant la méthode de Bonferroni et/ou le contrôle du taux de fausses découvertes (*false discovery rate*, FDR) (tableau III.C.3.1)

Analyses basées sur les voies métaboliques

Neuf études ont évalué les associations entre les expositions environnementales et les voies métaboliques (65,66,215,225,226,228,230–232) (tableau III.C.3.1). Certaines de ces études ont également appliqué une analyse topologique (66,228). Parmi les neuf études ayant analysé les voies métaboliques, toutes sauf une (226) ont appliqué une correction pour tests multiples (le plus souvent le contrôle du taux de fausses découvertes, *false discovery rate* – FDR) (tableau III.C.3.1).

(2) Méthodes d'analyses multidimensionnelles

Douze études ont appliqué des méthodes statistiques multidimensionnelles (65,66,212–214,220–223,228,229,232) (tableau III.C.3.1).

f) Représentations graphiques

Nous avons observé une hétérogénéité dans les représentations graphiques utilisées pour présenter les résultats (tableau III.C.3.1). Onze études ont eu recours à des *forest plots*

(212,214–216,220–223,226–228), dix à des *heatmaps* (65,215–217,220,222,228,230–232), et cinq à des *volcano plots* (66,218,221,230,232). Une large gamme d'autres représentations graphiques a également été utilisée dans les articles, notamment des diagrammes de Venn, des courbes ROC (*receiver operating characteristic*), des *biplots*, des *boxplots*, des histogrammes horizontaux, des *Manhattan plots* et des *bubble plots*.

g) Analyses statistiques selon la question scientifique

(1) Question scientifique relative aux processus ou voies métaboliques

Parmi les 13 études dont la question scientifique était classée comme « Quels processus ou voies métaboliques spécifiques sont influencés par l'exposition ? », 10 ont évalué une association avec des métabolites individuels en appliquant des régressions linéaires (64,216–218,226–228,230,231), ou une corrélation de Spearman (225) ou une analyse de covariance (226).

Par ailleurs, 4 études ont utilisé des analyses multidimensionnelles : l'une a utilisé une régression multiple multivariable (*multivariable multiple linear regression*, MMLR) (228), une autre une analyse en composantes principales (ACP) (212) et 3 ont utilisé un modèle elastic net (212,213,222). Enfin, une étude a utilisé un modèle de Cox, un modèle de Weibull et une corrélation partielle (226).

Quatre de ces 13 études ont évalué une association au niveau des voies métaboliques à l'aide de l'ORA (228,230,231) ou de la méthode de Fisher (225).

Les résultats sont présentés dans le Tableau III.C.3.1, le Tableau II.C.3.3 et le Tableau supplémentaire III.C.3.2.

(2) Question scientifique relative à la médiation

Parmi les 10 études ayant une question scientifique liée à la médiation, 8 ont utilisé des modèles de régression (65,66,215,219–221,223,224), 3 ont eu recours à des modèles d'elastic net et de Cox (65,222,223), une a utilisé l'UMAP (*uniform manifold approximation and projection*) (220), une a réalisé des analyses PLS ainsi qu'une analyse d'intégration N (66), une a utilisé la méthode de pénalisation Lasso (*least absolute shrinkage and selection operator*) (229), et une a effectué des analyses en composantes principales, de randomisation mendélienne et de dépendance directionnelle (221). Enfin, 3 études ont réalisé des analyses de voies métaboliques (65,66,215).

Les résultats sont présentés dans le Tableau III.C.3.1, le Tableau III.C.3.3 et le Tableau supplémentaire III.C.3.2.

h) Relation entre la question scientifique et l'hypothèse H1, ainsi qu'entre l'hypothèse H1 et le critère de jugement statistique

(1) Relation entre la question scientifique et l'hypothèse H1

Question scientifique liée aux processus métaboliques ou aux voies métaboliques

Pour une question scientifique liée aux processus ou aux voies métaboliques, les hypothèses H1 étaient considérées comme pertinentes lorsqu'une analyse de voies métaboliques était

mentionnée (Tableau III.C.3.3 et Tableau supplémentaire III.C.3.2). Cela concernait 3 des 13 études ayant abordé cette question scientifique (225,227,231).

L'alignement était considéré comme partiel lorsque des associations avec des métabolites individuels étaient testées, mais qu'aucune analyse au niveau des voies métaboliques n'était réalisée. Cela concernait 8 des 13 études ayant abordé cette question scientifique (64,212,213,216–218,222,230).

Pour les deux études dont l'objectif incluait le terme « profil », et pour lesquelles l'hypothèse H1 n'était pas clairement identifiable, l'alignement n'a pas été coté (226,228).

Question scientifique liée à la médiation

Les 10 études ayant une question scientifique liée à la médiation présentaient toutes une hypothèse H1 classée comme relevant de la médiation et donc pertinente (65,66,215,219–224,229) (tableau III.C.3.3 et tableau supplémentaire III.C.3.2).

Question scientifique supplémentaire (« Il existe une interaction avec l'IMC »)

Pour les 2 études ayant cette question scientifique (216,218), les deux présentaient une hypothèse H1 incluant une interaction et donc pertinente (Tableau III.C.3.3 et Tableau supplémentaire III.C.3.2).

Au total, quinze des 24 études incluses présentaient une hypothèse H1 classée comme pertinente au regard de la question scientifique (65,66,215,216,218–225,227,229,231).

(2) Relation entre l'hypothèse H1 et l'outcome statistique

Question scientifique liée aux processus ou aux voies métaboliques

Les outcomes statistiques étaient considérés comme pertinents lorsqu'ils correspondaient à l'exposition et à la variable métabolique spécifiées dans l'hypothèse H1. Lorsque la variable métabolique n'était pas précisément définie dans H1, nous avons classé les outcomes statistiques comme pertinents. Nous reconnaissons que cette situation peut relever d'un manque de clarté de la rédaction plutôt que d'une incohérence méthodologique.

Parmi les 9 études dont l'hypothèse H1 était « Il existe une association entre au moins un métabolite et une exposition environnementale » (64,212,213,216–218,222,225,230), ainsi que parmi les 3 études dont l'hypothèse H1 était « Il existe une association entre au moins une voie métabolique et une exposition environnementale » (225,227,231), toutes ont rapporté des outcomes statistiques pertinents au regard de cette hypothèse.

Parmi les 2 études présentant une hypothèse H1 imprécise, « Il existe une association entre au moins un ? et une exposition environnementale » (226,228), les outcomes statistiques étaient dans tous les cas pertinents au regard de l'hypothèse H1.

Par ailleurs, l'étude dont la question scientifique n'a pas pu être classée (232) comportait deux hypothèses H1 (« Il existe une association entre au moins un métabolite et une exposition environnementale » et « Il existe une association entre au moins une voie métabolique et une exposition environnementale »), et les outcomes statistiques étaient pertinents pour chacune des hypothèses H1. Cette étude aurait donc pu être classée dans cette catégorie de question scientifique sur la base de la méthodologie.

Les résultats sont présentés dans le tableau III.C.3.3 et le tableau supplémentaire III.C.3.2.

Question scientifique liée à la médiation

Les outcomes statistiques étaient considérés comme pertinents lorsqu'un effet de médiation était testé ou lorsque la proportion de l'effet médié était évaluée. L'alignement était considéré comme partiel lorsque les auteurs examinaient l'association entre l'exposition et les métabolites, puis l'association entre les métabolites liés à l'exposition et le phénotype, sans calculer la proportion médiée. L'alignement était considéré comme négatif lorsqu'aucun de ces deux éléments n'était abordé.

Parmi les 10 études dont l'hypothèse H1 était liée à la médiation, 5 études présentaient des outcomes statistiques pertinents (65,215,219,221,223), 4 étaient partiellement pertinentes (66,222,224,229), et 1 n'était pas pertinente (220). Les résultats sont présentés dans le tableau III.C.3.3 et le tableau supplémentaire III.C.3.2.

Au total, 19 des 24 études incluses présentaient un outcome statistique classé comme pertinent au regard de l'hypothèse H1 (64,65,212,213,215–219,221–223,225–228,230–232).

i) Validité statistique

Nous avons identifié plusieurs problèmes statistiques dans les publications examinées (tableau III.C.3.3).

Premièrement, l'analyse de voies métaboliques utilisait parfois la méthode de Fisher pour le test global. Cette méthode teste l'hypothèse H0 selon laquelle les p-values de chaque association entre un métabolite de la voie et l'exposition suivent une distribution uniforme (ce qui correspond à l'absence d'association entre l'exposition et la voie). Les deux publications ayant utilisé cette méthode (225,232) n'ont pas précisé comment celle-ci avait pu être adaptée. Cependant, la méthode de Fisher repose sur l'hypothèse que tous les tests individuels pour chaque métabolite sont indépendants, ce qui est peu probable pour des métabolites appartenant à une même voie métabolique (234). Par ailleurs, dans quatre études dont l'objectif correspondait à la médiation (65,215,219,223), le médiateur était (i) un score construit à partir de l'association entre l'exposition et les métabolites, avec une estimation de la médiation réalisée dans l'échantillon utilisé pour construire ce score, ou (ii) défini par la sélection des seuls métabolites significativement associés à l'exposition pour l'analyse de médiation, réalisée dans le même échantillon. Dans les deux cas, cela peut conduire à un surajustement (overfitting).

Enfin, une étude ne rapportait pas la méthode utilisée pour l'analyse de voies métaboliques (226).

Tableau III.C.3.1. Description des études

Auteur, année	Echantillon provenant de	Taille d'échantillon	Exposition	Méthodes statistiques :						
				Analyses de métabolite : modèles de base	Méthodes d'analyses de voies métaboliques	Analyses multidimensionnelles	Autres méthodes statistiques	Représentations graphiques des résultats	Correction des tests multiples (métabolites)	Correction des tests multiples (voies métaboliques)
Rhee J <i>et al.</i> , 2023	Prostate, Lung, Colorectal and Ovarian Cancer (PLCO)	3647	PFOS et PFOA	Régression linéaire	Aucune	Aucune	Analyse poolée	Scatter plots	Bonferroni	Non applicable
Cross AJ <i>et al.</i> , 2014	Prostate, Lung, Colorectal and Ovarian Cancer (PLCO)	509	Tabac	Régression linéaire, régression logistique	Aucune	Aucune	Aucune	Plot	Bonferroni	Non applicable
Gu F <i>et al.</i> , 2016	Environment And Genetics in Lung Cancer Etiology (EAGLE) ; Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) ; Alpha-Tocopherol, BetaCarotene Cancer Prevention (ATBC) ; et Shanghai Physical Activity Study (SPAS)	892	Tabac	Régression linéaire	Aucune	Aucune	Analyse poolée	Heatmap, forest plot	Bonferroni	Non applicable
Playdon MC <i>et al.</i> , 2017	Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC)	1336	Qualité de l'alimentation	Corrélation de Spearman	Méthode de Fisher	Aucune	Analyse poolée	Horizontal bar chart	Bonferroni	Bonferroni
Hernández-Alonso P <i>et al.</i> , 2019	PREDIMED	1157	Consommation de vin rouge	Aucune	Aucune	Elastic net (échantillons d'entraînement et de validation, validation croisée)	Aucune	Forest plot, courbe ROC	Non applicable	Non applicable
Hernández-Alonso P <i>et al.</i> , 2020	PREDIMED	1833	Apport en protéines	Aucune	Aucune	ACP, elastic net (échantillons d'entraînement et de validation, validation croisée)	Aucune	Forest plot, diagramme de Venn, biplot	Non applicable	Non applicable

Tableau III.C.3.1. Description des études (suite)

Auteur, année	Echantillon provenant de	Taille d'échantillon	Exposition	Méthodes statistiques :						
				Analyses de métabolite : modèles de base	Méthodes d'analyses de voies métaboliques	Analyses multidimensionnelles	Autres méthodes statistiques	Représentations graphiques des résultats	Correction des tests multiples (métabolites)	Correction des tests multiples (voies métaboliques)
Papandreou C <i>et al.</i> , 2019	PREDIMED	1664	Consommation de café	Aucune	Aucune	Elastic net (échantillons d'entraînement et de validation, validation croisée)	Méthode non paramétrique	Courbe ROC	Non applicable	Non applicable
Drouin-Chartier JP <i>et al.</i> , 2021	PREDIMED ; Nurses' Health Study (NHS) ; Nurses' Health Study II (NHSII) ; et Health Professionals Follow-Up Study (HPFS)	6765	Consommation de produits laitiers	Aucune	Aucune	Elastic net regression (échantillons d'entraînement et de validation, validation croisée)	Modèle de Cox	Heatmap, plot, forest plot	Non applicable	Non applicable
Wang F <i>et al.</i> , 2022	Nurses' Health Study (NHS) ; Nurses' Health Study II (NHSII) ; et Health Professionals Follow-Up Study (HPFS)	10684	Régime alimentaire à base de plantes	Régression linéaire	Aucune	Elastic net (échantillons d'entraînement et de validation, validation croisée)	Modèle de Cox	Pie chart, diagramme de Venn, correlation plot, plot, forest plot, horizontal bar chart, matrice de corrélation	Bonferroni	Non applicable
Tessier AJ <i>et al.</i> , 2024	Nurses' Health Study (NHS) ; Nurses' Health Study II (NHSII) ; Health Professionals Follow-Up Study (HPFS) ; et Women's Health Initiative (WHI)	13056	Mode de vie sain	Régression linéaire, régression logistique	MSEA	Elastic net regression (échantillons d'entraînement et de validation)	Modèle de Cox	Barplot, heatmap, boxplot	Bonferroni	FDR
Huang T <i>et al.</i> , 2022	Nurses' Health Study II (NHSII)	1014	Abus précoces	Régression linéaire	MSEA	Aucune	Analyse poolée, approche contrefactuelle	Forest plot, heatmap, correlation map	FDR	FDR
Tabung FK <i>et al.</i> , 2020	Women's Health initiative (WHI)	1919	Régime alimentaire	Régression linéaire	Aucune	Aucune	Aucune	Heatmap, forest plot	FDR	Non applicable
Che X <i>et al.</i> , 2024	Boston Birth Cohort (BBC)	1410	Adhésion au régime de type Méditerranéen	Régression linéaire	Aucune	Aucune	Aucune	Volcano plot	FDR	Non applicable

Tableau III.C.3.1. Description des études (suite)

Auteur, année	Echantillon provenant de	Taille d'échantillon	Exposition	Méthodes statistiques :						
				Analyses de métabolite : modèles de base	Méthodes d'analyses de voies métaboliques	Analyses multidimensionnelles	Autres méthodes statistiques	Représentations graphiques des résultats	Correction des tests multiples (métabolites)	Correction des tests multiples (voies métaboliques)
Choi JJ <i>et al.</i> , 2023	Wisconsin Registry for Alzheimer's Prevention (WRAP)	1266	Tabac	Régression (modèles linéaires mixtes)	Aucune	Aucune	Aucune	Correlation chart, Manhattan plot	FDR	Non applicable
Chen X <i>et al.</i> , 2022	Complex Lipids in Mothers and Babies (CLIMB) birth cohort	1020	Pollution de l'air	Régression logistique	Aucune	UMAP	Aucune	Heatmap, line plot, forest plot, circles plot	FDR	Non applicable
Salihovic S <i>et al.</i> , 2019	Prospective Investigation of the Vasculature in Uppsala Seniors study (PIVUS)	965	PFAS	Régression linéaire	ORA, analyse de topologie	MMLR	Aucune	Forest plot, bubble plot, heatmap	Bonferroni	FDR
Cohen CC <i>et al.</i> , 2022	Exploring Perinatal Outcomes Among Children study (EPOCH)	597	Consommations de boissons sucrées	Aucune	Aucune	LASSO (validation croisée, bootstrap)	Modèles d'équations structurelles	Scatter plot	Non applicable	Non applicable
Alcolea JA <i>et al.</i> , 2023	Multi Case-Control Spain (MCC-Spain) project	585	Trihalométhanes et nitrates dans l'eau potable	Régression linéaire	ORA, analyse de topologie	PLS et sPLS (validation croisée), PLS-DA (validation croisée), analyse par intégration N (échantillons d'entraînement et de validation)	Aucune	Upset plot, volcano plot, diagrammes de Venn, circos plot	FDR	FDR
Walker ME <i>et al.</i> , 2020	Framingham Offspring Study	2208	Profils alimentaires sains	Régression linéaire	ORA uniquement sur les protéines (pas sur les métabolites)	Aucune	Aucune	Diagramme de Venn, volcano plot, heatmap	Bonferroni et FDR	FDR

Tableau III.C.3.1. Description des études (fin)

Auteur, année	Echantillon provenant de	Taille d'échantillon	Exposition	Méthodes statistiques :						
				Analyses de métabolite : modèles de base	Méthodes d'analyses de voies métaboliques	Analyses multidimensionnelles	Autres méthodes statistiques	Représentations graphiques des résultats	Correction des tests multiples (métabolites)	Correction des tests multiples (voies métaboliques)
Zheng Y <i>et al.</i> , 2014 (a)	Atherosclerosis Risk in Communities (ARIC)	1977	Habitudes de consommation d'alcool	Régression linéaire, analyse de covariance	Non précisée	Aucune	Modèles de Weibull, modèles de Cox, corrélations partielles	Forest plot, curve plot	Bonferroni	Aucune
Zheng Y <i>et al.</i> , 2014 (b)	Atherosclerosis Risk in Communities (ARIC)	1977	Habitudes alimentaires	Régression linéaire	Aucune	Aucune	Aucune	Forest plot	Bonferroni	Non applicable
Beuchel C <i>et al.</i> , 2019	LIFE-adult ; LIFE-Heart et Sorbs	16222	Facteurs cliniques et liés au mode de vie	Régression linéaire	ORA	Aucune	Aucune	Plot, heatmap, network, boxplot, correlation map	Bonferroni	FDR
Wang Y <i>et al.</i> , 2020	China Health and Nutrition Survey (CHNS)	500	Lieu de résidence urbain ou rural	Régression linéaire	ORA, méthode de Fisher	Classification par forêts aléatoires (échantillons d'entraînement et de validation), O-PLS-DA (validation croisée, permutation), ACP	Aucune	Score plot, courbe ROC, volcano plot, heatmap, plot, boxplot	FDR	FDR
Zhang R <i>et al.</i> , 2022	Bogalusa Heart Study (BHS)	1252	Tabac	Régression linéaire	Aucune	ACP	Analyses de randomisation mendélienne, analyse de dépendance directionnelle	Volcano plot, forest plot	Bonferroni	Non applicable

Zheng Y *et al.*, 2014 (a) : Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study

Zheng Y *et al.*, 2014 (b) : Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study

Abréviations: PFOS = perfluorooctane sulfonate ; PFOA = perfluorooctanoic acid ; PFAS = perfluoroalkyl substances ; MSEA = metabolite set enrichment analysis ; ORA = over representation analysis ; ACP = analyse en composantes principales ; UMAP = uniform manifold approximation and projection ; MMLR = multivariable multiple regression ; LASSO = least absolute shrinkage and selection operator ; PLS = partial least squares ; sPLS = sparse partial least squares ; PLS-DA = partial least squares-discriminant analysis ; O-PLS-DA = orthogonal partial least squares-discriminant analysis ; ROC = receiver operating characteristic ; FDR = false discovery rate

Tableau III.C.3.2. Evaluation de la qualité des études et du risque de biais (score de ROBINS)

Auteur, année	Taille d'échantillon	Exposition	Score de ROBINS
Rhee J <i>et al.</i> , 2023	3647	PFOS et PFOA	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Cross AJ <i>et al.</i> , 2014	509	Tabac	Certaines préoccupations
Gu F <i>et al.</i> , 2016	892	Tabac	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Playdon MC <i>et al.</i> , 2017	1336	Qualité de l'alimentation	Certaines préoccupations
Hernández-Alonso P <i>et al.</i> , 2019	1157	Consommation de vin rouge	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Hernández-Alonso P <i>et al.</i> , 2020	1833	Apport en protéines	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Papandreou C <i>et al.</i> , 2019	1664	Consommation de café	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Drouin-Chartier JP <i>et al.</i> , 2021	6765	Consommation de produits laitiers	Certaines préoccupations
Wang F <i>et al.</i> , 2022	10684	Régime alimentaire à base de plantes	Certaines préoccupations
Tessier AJ <i>et al.</i> , 2024	13056	Mode de vie sain	Certaines préoccupations
Huang T <i>et al.</i> , 2022	1014	Abus précoces	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Tabung FK <i>et al.</i> , 2020	1919	Régime alimentaire	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Che X <i>et al.</i> , 2024	1410	Adhésion au régime de type Méditerranéen	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Choi JJ <i>et al.</i> , 2023	1266	Tabac	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Chen X <i>et al.</i> , 2022	1020	Pollution de l'air	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels
Salihovic S <i>et al.</i> , 2019	965	PFAS	Certaines préoccupations
Cohen CC <i>et al.</i> , 2022	597	Consommations de boissons sucrées	Certaines préoccupations
Alcolea JA <i>et al.</i> , 2023	585	Trihalométhanes et nitrates dans l'eau potable	Certaines préoccupations

Tableau III.C.3.2. Evaluation de la qualité des études et du risque de biais (score de ROBINS) (fin)

Auteur, année	Taille d'échantillon	Exposition	Score de ROBINS
Walker ME <i>et al.</i> , 2020	2208	Profils alimentaires sains	Certaines préoccupations
Zheng Y <i>et al.</i> , 2014 (a)	1977	Habitudes de consommation d'alcool	Certaines préoccupations
Zheng Y <i>et al.</i> , 2014 (b)	1977	Habitudes alimentaires	Certaines préoccupations
Beuchel C <i>et al.</i> , 2019	16222	Facteurs cliniques et liés au mode de vie	Certaines préoccupations
Wang Y <i>et al.</i> , 2020	500	Lieu de résidence urbain ou rural	Certaines préoccupations
Zhang R <i>et al.</i> , 2022	1252	Tabac	Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels

Zheng Y *et al.*, 2014 (a) : Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study

Zheng Y *et al.*, 2014 (b) : Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study

Abréviations : PFOS = perfluorooctane sulfonate ; PFOA = perfluorooctanoic acid ; PFAS = perfluoroalkyl substances

L'outil ROBINS-I comporte 5 catégories de risque de biais: Faible risque de biais ; Faible risque de biais, à l'exception de préoccupations concernant les facteurs de confusion résiduels ; Certaines préoccupations ; Risque élevé de biais ; Risque très élevé de biais

Tableau III.C.3.3. Cohérence entre la question scientifique, l'hypothèse H1, l'outcome statistique et l'absence de problèmes statistiques

Auteur, année	Question scientifique	Formulation des hypothèses H1	Méthodes d'analyse statistique	Outcome statistique	Cohérence		
					L'hypothèse H1 était appropriée à la question scientifique	L'outcome statistique était approprié à l'hypothèse H1	Absence de problèmes statistiques identifiés
Rhee J et al., 2023	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins un métabolite et une exposition environnementale	1) Régressions linéaires	a) Associations pour les métabolites individuels	A) Partiellement	A) Oui	Oui
Cross AJ et al., 2014	A) Identifier des proxys d'exposition B) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Non concerné par H1 B) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	1) Régressions linéaires 2) Régressions logistiques	a) Associations pour les métabolites individuels b) Associations entre les métabolites associés à l'exposition et le phénotype	A) Non concerné B) Oui	A) Non concerné B) Partiellement	Oui
Gu F et al., 2016	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins un métabolite et une exposition environnementale	1) Régressions linéaires	a) Associations pour les métabolites individuels	A) Partiellement	A) Oui	Oui
Playdon MC et al., 2017	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une association entre au moins une voie métabolique et une exposition environnementale	1) Corrélation de Spearman 2) Méthode de Fisher	a) Corrélations pour les métabolites individuels b) Associations par voie métabolique	A) Oui	A) Oui B) Oui	Non (les tests ne sont pas indépendants)

Tableau III.C.3.3. Cohérence entre la question scientifique, l'hypothèse H1, l'outcome statistique et l'absence de problèmes statistiques (suite)

Auteur, année	Question scientifique	Formulation des hypothèses H1	Méthodes d'analyse statistique	Outcome statistique	Cohérence		
					L'hypothèse H1 était appropriée à la question scientifique	L'outcome statistique était approprié à l'hypothèse H1	Absence de problèmes statistiques identifiés
Hernández-Alonso P et al., 2019	A) Non défini B) Identifier des proxies d'exposition	A) H1 non définie B) Non concerné par H1	1) Elastic net	a) Associations pour les métabolites individuels	A) Non défini B) Non concerné	A) Non concerné	Oui
Hernández-Alonso P et al., 2020	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins un métabolite et une exposition environnementale	1) Elastic net 2) ACP	a) Associations pour les métabolites individuels b) Profil métabolomique (matériel supplémentaire)	A) Partiellement	A) Oui	Oui
Papandreou C et al., 2019	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Identifier des proxies d'exposition	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Non concerné par H1	1) Elastic net 2) Courbe ROC	a) Associations pour les métabolites individuels b) Score pondéré	A) Partiellement B) Non concerné	A) Oui	Oui
Drouin-Chartier JP et al., 2021	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	1) Elastic net 2) Modèles de Cox	a) Associations pour les métabolites individuels b) Associations entre les métabolites associés à l'exposition et le phénotype	A) Partiellement B) Oui	A) Oui B) Partiellement	Oui

Tableau III.C.3.3. Cohérence entre la question scientifique, l'hypothèse H1, l'outcome statistique et l'absence de problèmes statistiques (suite)

Auteur, année	Question scientifique	Formulation des hypothèses H1	Méthodes d'analyse statistique	Outcome statistique	Cohérence		
					L'hypothèse H1 était appropriée à la question scientifique	L'outcome statistique était approprié à l'hypothèse H1	Absence de problèmes statistiques identifiés
Wang F et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype B) Au moins un score métabolique est un médiateur dans la relation entre l'exposition et le phénotype	1) Elastic net 2) Régressions linéaires 3) Modèles de Cox 4) Pourcentage de médiation	a) Associations pour les métabolites individuels b) Score pondéré, proportion médiée	A) Oui B) Oui	A) Oui B) Oui	Non (overfitting)
Tessier AJ et al., 2024	A) Non défini B) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) H1 non définie B) Au moins ? est un médiateur dans la relation entre l'exposition et le phénotype	1) Régressions linéaires et logistiques 2) MSEA 3) Elastic net 4) Modèles de Cox 5) Pourcentage de médiation	a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Score pondéré, associations avec le phénotype, proportion médiée	A) Non défini B) Oui	A) Non évalué B) Oui	Non (overfitting)
Huang T et al., 2022	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Evaluer l'interaction avec l'IMC	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une interaction avec l'IMC	1) Régressions linéaires 2) MSEA 3) Analyses en sous-groupes selon l'IMC	a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Associations pour les métabolites individuels, selon la classe d'IMC	A) Partiellement B) Oui	A) Oui B) Non	Oui
Tabung FK et al., 2020	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Au moins ? est un médiateur dans la relation entre l'exposition et le phénotype	1) Régressions linéaires 2) MSEA 3) Pourcentage de médiation	a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Score pondéré, proportion médiée	A) Oui	A) Oui	Non (overfitting)

Tableau III.C.3.3. Cohérence entre la question scientifique, l'hypothèse H1, l'outcome statistique et l'absence de problèmes statistiques (suite)

Auteur, année	Question scientifique	Formulation des hypothèses H1	Méthodes d'analyse statistique	Outcome statistique	Cohérence		
					L'hypothèse H1 était appropriée à la question scientifique	L'outcome statistique était approprié à l'hypothèse H1	Absence de problèmes statistiques identifiés
Che X et al., 2024	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Evaluer l'interaction avec les caractéristiques maternelles	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une interaction avec les caractéristiques maternelles	1) Régressions linéaires 2) Test d'interaction	a) Associations pour les métabolites individuels b) Interaction	A) Partiellement B) Oui	A) Oui B) Oui	Oui
Choi JJ et al., 2023	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	1) Modèles linéaires mixtes	a) Associations pour les métabolites individuels b) Proportion médiée	A) Oui	A) Oui	Non (overfitting)
Chen X et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Au moins ? est un médiateur dans la relation entre l'exposition et le phénotype	1) Régressions logistiques 2) UMAP	a) Associations pour les métabolites individuels b) Clusters	A) Oui	A) Non	Oui
Salihovic S et al., 2019	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins ? et une exposition environnementale	1) Régressions linéaires 2) ORA 3) MMLR	a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Test de différenciation pour les 6 profils	A) Non évalué	A) Oui	Oui

Tableau III.C.3.3. Cohérence entre la question scientifique, l'hypothèse H1, l'outcome statistique et l'absence de problèmes statistiques (suite)

Auteur, année	Question scientifique	Formulation des hypothèses H1	Méthodes d'analyse statistique	Outcome statistique	Cohérence		
					L'hypothèse H1 était appropriée à la question scientifique	L'outcome statistique était approprié à l'hypothèse H1	Absence de problèmes statistiques identifiés
Cohen CC et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	1) Modèles mixtes 2) Lasso	a) Associations entre l'exposition et le phénotype b) Associations entre l'exposition et les métabolites individuels c) Associations entre les métabolites associés à l'exposition et le phénotype	A) Oui	A) Partiellement	Oui
Alcolea JA et al., 2023	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype B) Au moins une voie métabolique est un médiateur dans la relation entre l'exposition et le phénotype	1) Régressions linéaires 2) ORA 3) PLS, sPLS, PLS-DA, analyse d'intégration N	a) Associations entre les métabolites individuels et l'exposition b) Associations entre les métabolites associés à l'exposition et le phénotype c) Associations entre les voies métaboliques et l'exposition d) Associations between exposure-associated pathways and phenotype	A) Oui	A) Partiellement	Oui
Walker ME et al., 2020	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins un métabolite et une exposition environnementale	1) Régressions linéaires	a) Associations pour les métabolites individuels	A) Partiellement	A) Oui	Oui
Zheng Y et al., 2014 (a)	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins ? et une exposition environnementale	1) Régressions linéaires, ANCOVA 2) Méthode non spécifiée pour les voies métaboliques 3) Modèles de Cox, de Weibull, corrélations partielles	a) Associations pour les métabolites individuels b) Score non pondéré, association entre le score et le phénotype	A) Non évalué	A) Oui	Non (méthode non spécifiée pour l'analyse des voies)

Tableau III.C.3.3. Cohérence entre la question scientifique, l'hypothèse H1, l'outcome statistique et l'absence de problèmes statistiques (fin)

Auteur, année	Question scientifique	Formulation des hypothèses H1	Méthodes d'analyse statistique	Outcome statistique	Cohérence		
					L'hypothèse H1 était appropriée à la question scientifique	L'outcome statistique était approprié à l'hypothèse H1	Absence de problèmes statistiques identifiés
Zheng Y et al., 2014 (b)	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins une voie métabolique et une exposition environnementale	1) Régressions linéaires	a) Associations pour les métabolites individuels	A) Oui	A) Non	Oui
Beuchel C et al., 2019	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	A) Il existe une association entre au moins une voie métabolique et une exposition environnementale	1) Régressions linéaires 2) ORA	a) Associations pour les métabolites individuels b) Associations par voie métabolique	A) Oui	A) Oui	Oui
Wang Y et al., 2020	Non classifiable	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une association entre au moins une voie métabolique et une exposition environnementale	1) Régressions linéaires 2) ORA et Méthode de Fisher 3) Classification par forêt aléatoire, O-PLS-DA, ACP	a) Séparation selon le statut rural/urbain b) Associations pour les métabolites individuels c) Associations par voie métabolique	A) Non évalué	A) Oui	Non (les tests ne sont pas indépendants)
Zhang R et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	1) Régressions linéaires 2) Randomisation mendélienne et analyse de dépendance directionnelle 3) ACP (effet médiateur des six premières composantes principales)	a) Associations pour les métabolites individuels b) Proportion médiée	A) Oui	A) Oui	Oui

Zheng Y et al., 2014 (a) : Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study

Zheng Y et al., 2014 (b): Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study

Abréviations : MSEA: metabolite set enrichment analysis ; ORA: over representation analysis ; ACP : analyse en composantes principales ; UMAP : uniform manifold approximation and projection ; MMLR : multivariable multiple regression ; Lasso : least absolute shrinkage and selection operator ; PLS : partial least squares ; sPLS : sparse partial least squares ; PLS-DA : partial least squares discriminant analysis ; O PLS-DA : orthogonal partial least squares discriminant analysis

Tableau supplémentaire III.C.3.1. Descriptions détaillées des populations d'étude

Auteur, année	Pays	Période d'étude	Type d'étude	Population				
				taille d'échantillon	méthode de recrutement	âge	sexe	caractéristiques (population générale ou population spécifique)
Rhee J <i>et al.</i> , 2023	États-Unis	1993-2001	Cas-témoins imbriquée	3647	Participants à huit études cas-témoins imbriquées de l'étude PLCO	Médiane [IIQ], années : 63 [60-67]	Femmes, % : 51	Cancer (cas) et témoins
Cross AJ <i>et al.</i> , 2014	États-Unis	1993-2001	Cas-témoins imbriquée	509	Issus de l'étude PLCO	Moyenne (ET), années : Cas = 64,3 (5,1) ; Contrôles = 64,3 (5,1)	Femmes, N (%) : Cas = 112 (43,92) ; Contrôles = 111 (43,7)	Cancer (cas) et témoins
Gu F <i>et al.</i> , 2016	Italie, États-Unis, Chine et Finlande	EAGLE : 2002-2005 ; PLCO : 1993-2001 ; ATBC : 1985-1988 ; SPAS : 2007-2008	Transversale	892	Issus des études EAGLE, PLCO, ATBC et SPAS	Moyenne (ET), années : EAGLE = 62,3 (9,2) ; PLCO = 64,3 (5,1) ; ATBC = 57,7 (5,0) ; SPAS = 54,9 (9,2)	Femmes, N (%) : EAGLE = 20 (20) ; PLCO = 223 (43,8) ; ATBC = 0 ; SPAS = 0	Population générale
Playdon MC <i>et al.</i> , 2017	Finlande	2013-2015	Transversale (provenant de 5 études cas-témoins)	1336	Issus de l'étude ATBC	Médiane [IIQ], années : 57 [54-62]	Hommes, N (%) : 1336 (100)	Hommes fumeurs
Hernández-Alonso P <i>et al.</i> , 2019	Espagne	PREDIMED : 2003-2009	Transversale	1157	Issus de l'étude PREDIMED	Moyenne (ET), années : 67,69 (6,04)	Hommes, N (%) : 436 (37,7)	Population à haut risque de maladies cardiovasculaires et de diabète
Hernández-Alonso P <i>et al.</i> , 2020	Espagne	PREDIMED : 2003-2009	Transversale	1833	Issus de l'étude PREDIMED	Médiane [IIQ], années = 67 [62-72]	Femmes, N (%) : 1055 (57,6)	Population à haut risque de maladies cardiovasculaires et de diabète

Tableau supplémentaire III.C.3.1. Descriptions détaillées des populations d'étude (suite)

Auteur, année	Pays	Période d'étude	Type d'étude	Population				
				taille d'échantillon	méthode de recrutement	âge	sexe	caractéristiques (population générale ou population spécifique)
Papandreou C <i>et al.</i> , 2019	Espagne	PREDIMED : 2003-2009	Transversale	1664	Issus de l'étude PREDIMED	Moyenne (ET), années = 67,14 (6)	Hommes, N (%) : 693 (41,6)	Population à haut risque de maladies cardiovasculaires et de diabète
Drouin-Chartier JP <i>et al.</i> , 2021	Espagne et États-Unis	PREDIMED : 2003-2009 ; NHS : 1989-1990 ; NHSII : 1996-1999 ; HPFS : 1993-1998	Multi-cohorte	6765	Population de découverte = étude PREDIMED à l'inclusion ; Population de confirmation = PREDIMED à 1 an de suivi et 3 cohortes américaines: NHS, NHSII, et HPFS	Moyenne (ET), années : PREDIMED = 67,2 (6,1) ; PREDIMED année 1 = 68,2 (6,0) ; Cohortes américaines = 56,0 (9,6)	Femmes, N (%) : PREDIMED = 1055 (57,6) ; PREDIMED à 1 an = 875 (57,5) ; Cohortes américaines = 3747 (76,0)	Population à haut risque de maladies cardiovasculaires (PREDIMED), infirmiers (NHS and NHSII), et professionnels de santé masculins (HPFS)
Wang F <i>et al.</i> , 2022	États-Unis	NHS : entre 1989 et 1990 ; NHSII : entre 1996 et 1999 ; HPFS : entre 1993 et 1995	Cohorte	10864	Issus de 3 cohortes américaines: NHS, NHSII et HPFS	Moyenne (ET), années : 54 (10)	Femmes, % : 83	Infirmiers (NHS and NHSII), professionnels de santé masculins (HPFS)
Tessier AJ <i>et al.</i> , 2024	États-Unis	NHS : entre 1989 et 1990 ; NHSII : entre 1996 et 1999 ; HPFS : entre 1993 et 1995 ; Women's Health Initiative (WHI) : en 1993	Cohorte	13056	Issus de 4 cohortes américaines: 3 cohortes de découvertes: NHS, NHSII, et HPFS ; et une cohorte de réplication externe (participants de l'étude WHI)	Moyenne (ET), années : 54,3 (9,0)	Femmes, N (%) : 9852 (85,8)	Infirmiers (NHS and NHSII), professionnels de santé masculins (HPFS), et femmes en postménopause (WHI)
Huang T <i>et al.</i> , 2022	États-Unis	1996-1999 (échantillon 1) ; 2013 (échantillon 2)	Cohorte	1014	2 sous-échantillons de l'étude NHSII	Description selon la sévérité des abus précoces : Moyenne (ET), années. Echantillon 1 : aucun = 44,2 (4,7) ; légère = 44,5 (4,7) ; modérée = 44,4 (4,8) ; sévère = 44,9 (4,8). Echantillon 2 : aucun = 60,4 (4,3) ; légère = 61,4 (4,6) ; modérée = 60,8 (3,6) ; sévère = 60,2 (4,4)	Femmes, N (%) : Echantillon 1 = 803 (100) ; Echantillon 2 = 211 (100)	Femmes infirmières
Tabung FK <i>et al.</i> , 2020	États-Unis	WHI : entre 1993 et 1998	Transversale	1919	Issus de l'étude WHI	Moyenne (ET), années : Echantillon de découverte : quantile 1 = 66,6 (7,4) ; quantile 2 = 67,4 (6,5) ; quantile 3 = 67,5 (6,8) ; quantile 4 = 67,1 (6,4) ; quantile 5 = 65,9 (7,6). Echantillon de réplication : quantile 1 = 67,2 (6,6) ; quantile 2 = 68,6 (6,1) ; quantile 3 = 68,2 (6,4) ; quantile 4 = 67,8 (6,4) ; quantile 5 = 66,0 (7,8)	Femmes, N (%) : 1919 (100)	Femmes en postménopause
Che X <i>et al.</i> , 2024	États-Unis	1998	Longitudinale	1410	Issus de la cohorte Boston Birth	Médiane [IIQ], années : 28,17 [22,89-33,54]	Femmes, N (%) : 1410 (100)	Femmes enceintes

Tableau supplémentaire III.C.3.1. Descriptions détaillées des populations d'étude (suite)

Auteur, année	Pays	Période d'étude	Type d'étude	Population				
				taille d'échantillon	méthode de recrutement	âge	sexe	caractéristiques (population générale ou population spécifique)
Choi JJ <i>et al.</i> , 2023	États-Unis	2020	Transversale	1266	Issus de l'étude WRAP: visite 2 à visite 6	Moyenne (ET), années : 58,5 (6,47)	Femmes, N (%) : 891 (70,4)	Population générale
Chen X <i>et al.</i> , 2022	Chine	2015-2016, entre 11 et 14 de semaines de grossesse	Longitudinale	1020	Issus de la cohorte de naissance CLIMB	Moyenne (ET), années : Dans le groupe du diabète gestationnel = 29 (27,32) ; Dans le groupe sans diabète gestationnel = 28	Femmes, N (%) : 1020 (100)	Femmes enceintes dans la population générale
Salihovic S <i>et al.</i> , 2019	Suède	2001-2004	Transversale	965	Issus de l'étude PIVUS: lettres d'invitation envoyées dans les 2 mois suivant le 70 ^e anniversaire	Moyenne (ET), années = 70,2 (0,2)	Femmes, % : 50,2	Population générale (autour du 70 ^e anniversaire)
Cohen CC <i>et al.</i> , 2022	États-Unis	2006-2009	Longitudinale	597	Issus de l'étude EPOCH	Moyenne (ET), années : Quartile 1 = 10,2 (1,5) ; Quartile 2 = 10,3 (1,4) ; Quartile 3 = 10,5 (1,4) ; Quartile 4 = 10,7 (1,5)	Hommes, N (%) : Quartile 1 = 63 (43) ; Quartile 2 = 72 (48) ; Quartile 3 = 75 (50) ; Quartile 4 = 87 (58)	Enfants
Alcolea JA <i>et al.</i> , 2023	Espagne	2007-2013	Cas-témoins	585	Issus de l'étude MCC-Spain	Moyenne (ET), années : Contrôles = 67,0 (7,7) ; Cas = 68,5 (9,2)	Hommes, N (%) : Contrôles = 182 (6,1) ; Cas = 191 (65,4)	Cancer colorectal (cas) et témoins

Tableau supplémentaire III.C.3.1. Descriptions détaillées des populations d'étude (fin)

Auteur, année	Pays	Période d'étude	Type d'étude	Population				
				taille d'échantillon	méthode de recrutement	âge	sexe	caractéristiques (population générale ou population spécifique)
Walker ME <i>et al.</i> , 2020	États-Unis	1991-1995	Transversale	2208	Issus de l'étude Framingham Offspring	Moyenne, années : 55	Femmes, % : 52	Population générale
Zheng Y <i>et al.</i> , 2014 (a)	États-Unis	1987-1989	Transversale	1977	Issus du centre d'investigation de Jackson de l'étude ARIC	Moyenne (SE), années : Echantillon de découverte = 52,90 (5,82) ; Echantillon de réplication = 52,83 (5,6)	Hommes, N (%) : Echantillon de découverte = 543 (36,2) ; Echantillon de réplication = 159 (33,3)	Afro-Américains
Zheng Y <i>et al.</i> , 2014 (b)	États-Unis	ARIC : 1987-1989	Transversale	1977	Issus de l'étude ARIC	Moyenne (SE), années : Echantillon de découverte = 52,92 (5,8) ; Echantillon de réplication = 52,74 (5,7)	Hommes, N (%) : Echantillon de découverte = 524 (34,9) ; Echantillon de réplication = 178 (37,3)	Afro-Américains
Beuchel C <i>et al.</i> , 2019	Allemagne	Plus de 10 ans	Cohorte	16222	Issus de 3 cohortes: LIFE-adult, LIFE-Heart et Sorbs	Médiane [IIQ], années : LIFE-adult = 57,91 [47,7-68,2] ; LIFE-Heart = 63,11 [54,4-71,7] ; Sorbs = 48,7 [35,6-60,9]	Femmes, N (%) : LIFE-adult = 4952 (52,2) ; LIFE-Heart = 1712 (29,7) ; Sorbs = 574 (58,9)	Individus phénotypés pour plusieurs maladies liées au mode de vie et les facteurs de risque associés correspondants (LIFE-adult), patients présentant une suspicion ou un diagnostic confirmé de coronaropathie recrutés au sein du Heart Center (LIFE-Heart), et une minorité ethnique d'origine slave résidant dans la région de Haute-Lusace, à l'est de la Saxe (Sorbs)
Wang Y <i>et al.</i> , 2020	Chine	2015	Transversale	500	Sous-échantillon de l'étude CHNS	Médiane [IIQ], années : Urbain = 54,6 [46,4-60,3] ; Rural = 51,8 [44,3-60,1]	Femmes, % : Urbain = 57,9 ; Rural = 60,4	Population générale
Zhang R <i>et al.</i> , 2022	États-Unis	Depuis 1973, enquête tous les 3 à 4 ans de l'enfance à l'âge adulte. Ici, cycle de visites 2013-2016	Transversale	1252	Issus de l'étude Bogalusa Health	Moyenne (ET), années : 48,2 (5,3)	Femmes, N (%) : 735 (58,7)	Afro-Américains et Blancs

Zheng Y *et al.*, 2014 (a) : Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study
 Zheng Y *et al.*, 2014 (b) : Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study

Abréviations: IIQ = intervalle interquartile ; ET = écart type

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Rhee J et al., 2023	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	"However, it is unclear <u>what mechanisms underlie these associations</u> because the biologic effects of these chemicals in humans remain poorly understood" + "To address this <u>knowledge gap</u> , we conducted a serum metabolomics study of circulating PFOS and PFOA "	A) Il existe une association entre au moins un métabolite et une exposition environnementale	PFOS et PFOA	a) Associations pour les métabolites individuels	1) Régressions linéaires
Cross AJ et al., 2014	A) Identifier des proxies d'exposition B) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	Our aim was to identify <u>metabolites that represent smoking habits</u> and individual variation in tobacco metabolism, and <u>investigate their association with colorectal cancer</u>	A) Non concerné par H1 B) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	Tabac	a) Associations pour les métabolites individuels b) Associations entre les métabolites associés à l'exposition et le phénotype	1) Régressions linéaires 2) Régressions logistiques
Gu F et al., 2016	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Identifying circulating metabolites related to cigarette smoking may provide <u>insight into the biological mechanisms</u> of smoking-related diseases and the nature of addiction	A) Il existe une association entre au moins un métabolite et une exposition environnementale	Tabac	a) Associations pour les métabolites individuels	1) Régressions linéaires

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique (suite)

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Playdon MC et al., 2017	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Our 2 purposes were as follows: 1) to <u>identify metabolites correlated</u> with 4 dietary patterns and their components [with the HEI-2010, the Alternate Mediterranean Diet Score (aMED), the HDI, and the Baltic Sea Diet (BSD)], and 2) to <u>gain insights into the potential biological mechanisms</u> influenced by diet quality	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une association entre au moins une voie métabolique et une exposition environnementale	Qualité de l'alimentation	a) Corrélations pour les métabolites individuels b) Associations par voie métabolique	1) Corrélation de Spearman 2) Méthode de Fisher
Hernández-Alonso P et al., 2019	A) Non défini B) Identifier des proxies d'exposition	We aimed to assess the systemic metabolomic profiles in relation to frequent RW consumption as well as the ability of a set of metabolites to <u>discriminate RW consumers</u>	A) H1 non définie B) Non concerné par H1	Consommation de vin rouge	a) Associations pour les métabolites individuels	1) Elastic net
Hernández-Alonso P et al., 2020	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Therefore, the main aim of the present study was to describe the set of metabolites associated with the intake of TP, animal protein (AP), plant protein (PP), and plant-to-animal protein ratio (PR), which could help us to <u>understand in the future the relationship between diet and cardiometabolic health</u>	A) Il existe une association entre au moins un métabolite et une exposition environnementale	Apport en protéines	a) Associations pour les métabolites individuels b) Profil métabolomique (matériel supplémentaire)	1) Elastic net 2) ACP

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique (suite)

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Papandreou C et al., 2019	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Identifier des proxies d'exposition	The aim of this study was to investigate associations of plasma metabolites with total, caffeinated, and decaffeinated coffee consumption. We also assessed the ability of metabolites to <u>discriminate between coffee consumption categories</u> . Metabolomics has the potential to advance nutritional epidemiology by objectively measuring metabolic products of foods, and might therefore better accurately reflect food exposure A comprehensive metabolite profiling may also provide a deeper <u>understanding of metabolic response foods providing new functional insight to their role in health</u>	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Non concerné par H1	Consommation de café	a) Associations pour les métabolites individuels b) Score pondéré	1) Elastic net 2) Courbe ROC
Drouin-Chartier JP et al., 2021	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	We aimed to identify the plasma <u>metabolites associated with total and specific dairy consumption</u> , and to evaluate <u>the association between the identified multi-metabolite profiles and T2D</u> . Overall, the <u>biological mechanisms underlying the complex relationship between dairy consumption and T2D risks remain poorly understood</u> . The integration of metabolomics into nutritional studies may provide valuable information on <u>the biochemical pathways underlying the relationship between dairy consumption and TD risk</u>	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	Consommation de produits laitiers	a) Associations pour les métabolites individuels b) Associations entre les métabolites associés à l'exposition et le phénotype	1) Elastic net 2) Modèles de Cox
Wang F et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	Plant-based diets, especially when rich in healthy plant foods, have been associated with a lower risk of type 2 diabetes. However, <u>whether plasma metabolite profiles related to plant-based diets reflect this association was unknown</u> Identifying metabolites associated with plant-based diets may provide new insights into the biological processes behind the association between plant-based diets and reduction in type 2 diabetes risk	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype B) Au moins un score métabolique est un médiateur dans la relation entre l'exposition et le phénotype	Régime alimentaire à base de plantes	a) Associations pour les métabolites individuels b) Score pondéré, proportion médiée	1) Elastic net 2) Régressions linéaires 3) Modèles de Cox 4) Pourcentage de médiation

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique (suite)

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Tessier AJ et al., 2024	A) Non défini B) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	However, the metabolic pathways of a healthy lifestyle and how they relate to mortality and longevity are unclear. We aimed to identify and replicate a healthy lifestyle metabolomic signature and examine how it is related to total and cause-specific mortality risk and longevity. Advances in metabolomic techniques hold promise in providing insights into underlying biological mechanisms, identifying high-risk individuals, and developing targeted preventive lifestyle interventions to reduce premature mortality. <u>We aimed to identify a metabolomic signature of a combined healthy lifestyle score in mid-life and examine its prospective association with all-cause and cause-specific mortality and longevity over up to 28 years of follow-up</u>	A) H1 non définie B) Au moins ? est un médiateur dans la relation entre l'exposition et le phénotype	Mode de vie sain	a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Score pondéré, associations avec le phénotype, proportion médiée	1) Régressions linéaires et logistiques 2) MSEA 3) Elastic net 4) Modèles de Cox 5) Pourcentage de médiation
Huang T et al., 2022	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Evaluer l'interaction avec l'IMC	Given the influence of dietary patterns on chronic systemic inflammation, dietary patterns associated with inflammation may therefore influence chronic disease outcomes. Elucidating the metabolomic profiles of inflammatory diets may therefore provide <u>insights into specific biological pathways</u> + We aimed to identify metabolites associated with the <u>inflammatory potential of diet and to examine the role of body weight in these associations</u>	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une interaction avec l'IMC	Abus précoce Régime alimentaire	a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Associations pour les métabolites individuels, selon la classe d'IMC	1) Régressions linéaires 2) MSEA 3) Analyses en sous-groupes selon l'IMC
Tabung FK et al., 2020	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	Metabolomic profiling may provide insights into biological mechanisms underlying the strong epidemiologic links observed between early abuse and cardiometabolic disorders in later life. We examined the associations between early abuse and plasma metabolomics in two non-overlapping subsamples of middle-aged women in the Nurses' Health Study II (NHSII). Prior work in this cohort has found early abuse, as well as psychosocial distress and unhealthy lifestyle during adulthood, was associated with increased cardiometabolic disease risk. We compared the metabolite association patterns for physical versus sexual abuse and for abuse occurring in childhood versus adolescence. We further evaluated whether and to what degree psychosocial and lifestyle factors during adulthood <u>mediated the association between early abuse and metabolic profiles at midlife</u>	A) Au moins ? est un médiateur dans la relation entre l'exposition et le phénotype		a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Score pondéré, proportion médiée	1) Régressions linéaires 2) MSEA 3) Pourcentage de médiation

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique (suite)

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Che X et al., 2024	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? B) Evaluer l'interaction avec les caractéristiques maternelles	Investigation into the consumption of an MSD and related postpartum metabolites, such as triacylglycerols (TAGs) and amino acids and adenosine, would <u>contribute to understanding how an MSD improves females' long-term health</u> . We aimed to identify postpartum metabolites associated with adherence to an MSD during pregnancy and to further <u>test whether these identified metabolites may vary by maternal characteristics</u> such as race and ethnicity, overweight or obesity (OWO), and hypertensive disorders of pregnancy (HDP)	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une interaction avec les caractéristiques maternelles	Adhésion au régime de type Méditerranéen	a) Associations pour les métabolites individuels b) Interaction	1) Régressions linéaires 2) Test d'interaction
Choi JJ et al., 2023	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	However, <u>the mediating role</u> that metabolites may play in the association between smoking and cognitive function has not been investigated. We sought to address this gap using <u>mediation analysis to identify whether metabolites</u> profiled from untargeted metabolomics in plasma and cerebrospinal fluid (CSF) are in the biological pathway between smoking and cognitive function. Identifying such metabolites may provide a better understanding of the mechanism linking smoking to cognitive decline	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	Tabac	a) Associations pour les métabolites individuels b) Proportion médiée	1) Modèles linéaires mixtes
Chen X et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	The aim of this research was to comprehensively assess <u>the relationship between air pollution, the maternal hair metabolome and GDM status</u> from the pre-conception period through the three trimesters of pregnancy	A) Au moins ? est un médiateur dans la relation entre l'exposition et le phénotype	Pollution de l'air	a) Associations pour les métabolites individuels b) Clusters	1) Régressions logistiques 2) UMAP

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique (suite)

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Salihovic S et al., 2019	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	From this perspective, metabolomics has emerged as a new tool in biomarker research. Metabolomics or metabolomic profiling encompasses the analysis of small molecules or metabolites (<1.5 kDa) in a sample and provides an unprecedented range of coverage when compared to traditional clinical methods. Non-targeted metabolomics has previously been used <u>to gain more in-depth knowledge about the biochemical pathways</u> associated with exposure to PFASs in an experimental setting; <u>yet little is known about the associations of PFASs and metabolomic profiles in humans</u> . Therefore, the main aim of this study was to investigate the associations between circulating levels of PFASs and the metabolome in a population-based sample of 965 individuals by using a non-targeted metabolomics approach. Because the chemical class of PFASs consists of compounds with differential properties and toxicities, we also sought to investigate whether the metabolic profiles were different between the six PFASs evaluated	A) Il existe une association entre au moins ? et une exposition environnementale	PFAS	a) Associations pour les métabolites individuels b) Associations par voie métabolique c) Test de différenciation pour les 6 profils	1) Régressions linéaires 2) ORA 3) MMLR
Cohen CC et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	The objective of this study was <u>to assess intermediary metabolic alterations that link sugar-sweetened beverage (SSB) intake to cardiometabolic (CM) risk factors in youth</u>	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	Consommation de boissons sucrées	a) Associations entre l'exposition et le phénotype b) Associations entre l'exposition et les métabolites individuels c) Associations entre les métabolites associés à l'exposition et le phénotype	1) Modèles mixtes 2) Lasso
Alcolea JA et al., 2023	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	Although evidence suggests that exposure to these widespread contaminants in drinking water may increase colorectal cancer risk, causal inference cannot be drawn due to the lack of understanding of the underlying biological mechanisms. Metabolomics offers a promising approach to gain insights into the relevant metabolic and molecular pathways involved. Some metabolites may be intermediate biomarkers, which directly reflect the underlying biochemical activity. Previous studies have linked THM exposure in swimming pools with changes in serum metabolomic signatures. However, the metabolomic profile associated with THMs and nitrate exposure in drinking water has not been evaluated so far. In light of this knowledge gap, we conducted an untargeted metabolomic study in the framework of the Multi Case-Control Spain project (MCC-Spain, www.mccspain.org). <u>We aimed to identify circulating metabolites associated with THMs and nitrate exposure, colorectal cancer, and the pathway between exposure and colorectal cancer</u>	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype B) Au moins une voie métabolique est un médiateur dans la relation entre l'exposition et le phénotype	Trihalométhanes et nitrates dans l'eau potable	a) Associations entre les métabolites individuels et l'exposition b) Associations entre les métabolites associés à l'exposition et le phénotype c) Associations entre les voies métaboliques et l'exposition d) Associations between exposure-associated pathways and phenotype	1) Régressions linéaires 2) ORA 3) PLS, sPLS, PLS-DA, analyse d'intégration N

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique (suite)

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Walker ME et al., 2020	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Multi-level omics analyses that combine high-dimensional molecular data from high throughput platforms can <u>provide a comprehensive assessment of intermediate phenotypes</u> (e.g., molecular endophenotypes) <u>that may help link diet (or other exposures) to more distal chronic disease phenotypes</u> . We hypothesized that proteomic and metabolic signatures of habitual dietary patterns encompass the respective functional states and metabolic consequences of unique dietary patterns including the sequelae of chronic disease. The objective of the present investigation was to determine the associations of the Alternative Healthy Eating Index (AHEI), Dietary Approaches to Stop Hypertension (DASH) diet, and a Mediterranean-style diet (MDS) with 1373 plasma proteins and 216 circulating metabolites in a sample of community-dwelling middle-aged adults	A) Il existe une association entre au moins un métabolite et une exposition environnementale	Profils alimentaires sains	a) Associations pour les métabolites individuels	1) Régressions linéaires
Zheng Y et al., 2014 (a)	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Metabolomics hold promise for novel biomarker discovery and development of an improved <u>understanding of underlying exposure-disease relations</u> . Metabolomic approaches have rarely been used in alcohol-related research. Therefore, we explored cross-sectional associations of multiple named metabolites quantified by an untargeted high-throughput mass spectrometry-based protocol with alcohol consumption habits and, in addition, longitudinal associations of selected alcohol-related metabolites <u>with future cardiovascular diseases</u> in a well-characterized, population-based sample of African Americans from the Atherosclerosis Risk in Communities (ARIC) Study	A) Il existe une association entre au moins ? et une exposition environnementale	Habitudes de consommation d'alcool	a) Associations pour les métabolites individuels b) Score non pondéré, association entre le score et le phénotype	1) Régressions linéaires, ANCOVA 2) Méthode non spécifiée pour les voies métaboliques 3) Modèles de Cox, de Weibull, corrélations partielles
Zheng Y et al., 2014 (b)	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Metabolomics could play an important role in the identification of novel biomarkers of dietary intake, as well as in the investigation of the mechanism of dietary effects on health and disease. In a well-characterized, population-based sample of African Americans from the Atherosclerosis Risk in Communities (ARIC) Study, we explored cross-sectional associations of multiple named metabolites quantified by an untargeted high-throughput chromatography/mass spectrometry-based protocol with usual dietary intake measured by a semiquantitative food frequency questionnaire (FFQ). <u>We aimed to identify metabolites pointing to pathways of nutrient metabolism</u>	A) Il existe une association entre au moins une voie métabolique et une exposition environnementale	Habitudes alimentaires	a) Associations pour les métabolites individuels	1) Régressions linéaires

Tableau supplémentaire 2. Cohérence entre la question scientifique, H1 et l'outcome statistique (fin)

Auteur, année	Question scientifique	Citation dans le texte permettant de classer l'objectif dans les catégories	Formulation des hypothèses H1	Exposition	Outcome statistique	Méthodes d'analyse statistique
Beuchel C et al., 2019	A) Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Targeted, high-throughput metabolomics using liquid chromatography-mass spectrometry (MS) increasingly gains momentum in epidemiology. Important fields of investigations are the understanding of the molecular basis of metabolism-related phenotypes and diseases and studying biomarkers for diagnostic and prognostic purposes. Furthermore, analysis of metabolomic features in relation to other molecular-genetic functional layers of the organism, e.g. genomics and transcriptomics, is <u>a promising approach to extend our knowledge of regulatory pathways and associated pathomechanisms</u>	A) Il existe une association entre au moins une voie métabolique et une exposition environnementale	Facteurs cliniques et liés au mode de vie	a) Associations pour les métabolites individuels b) Associations par voie métabolique	1) Régressions linéaires 2) ORA
Wang Y et al., 2020	Non classifiable	Not element found	A) Il existe une association entre au moins un métabolite et une exposition environnementale B) Il existe une association entre au moins une voie métabolique et une exposition environnementale	Lieu de résidence urbain ou rural	a) Séparation selon le statut rural/urbain b) Associations pour les métabolites individuels c) Associations par voie métabolique	1) Régressions linéaires 2) ORA et Méthode de Fisher 3) Classification par forêt aléatoire, O-PLS-DA, ACP
Zhang R et al., 2022	A) Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	Metabolomics can detect subtle changes in molecular profiles related to endogenous pathological processes, as well as dietary, environmental, and gut microbial influences. As the endpoint of internal and external molecular processes, examination of the human metabolome provides an opportunity to discover biological pathways linking cigarette smoking to weight loss. Previous studies have identified many circulating metabolites that were associated with cigarette smoking and/or body mass index (BMI), however, very few studies investigated <u>the mediating roles of metabolites in the effect of cigarette smoking on body weight</u> . Therefore, <u>we conducted an untargeted metabolomics study to identify serum metabolites that mediate the association between smoking and BMI among participants of The Bogalusa Heart Study</u>	A) Au moins un métabolite est un médiateur dans la relation entre l'exposition et le phénotype	Tabac	a) Associations pour les métabolites individuels b) Proportion médiée	1) Régressions linéaires 2) Randomisation mendélienne et analyse de dépendance directionnelle 3) ACP (effet médiateur des six premières composantes principales)

Zheng Y et al., 2014 (a) : Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study

Zheng Y et al., 2014 (b) : Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study

Abréviations : PFOS : perfluorooctane sulfonate ; PFOA : perfluorooctanoic acid ; PFAS : perfluoroalkyl substances ; MSEA : metabolite set enrichment analysis ; ORA : over representation analysis ; ACP : analyse en composantes principales ; UMAP : uniform manifold approximation and préjection ; MMLR : multivariable multiple regression ; Lasso : least absolute shrinkage and selection operator ; PLS : partial least squares ; sPLS : sparse partial least squares ; PLS-DA : partial least squares discriminant analysis ; O PLS-DA : orthogonal partial least squares discriminant analysis

Tableau supplémentaire III.C.3.3. Description des expositions et des mesures métaboliques

Auteur, année	Exposition			Métabolomique		
	types d'exposition	quantification	durée d'exposition	fluide	nombre de métabolites mesurés	nombre de métabolites annotés
Rhee J <i>et al.</i> , 2023	PFOS et PFOA	Non précisé	Non précisé	Sérum	Non précisé	1409
Cross AJ <i>et al.</i> , 2014	Tabac	Statut tabagisme (N) : Cas: actuels = 26 ; anciens = 120 ; jamais = 109. Contrôles : actuels = 19 ; anciens = 121 ; jamais = 114	Pendant 6 mois ou plus	Sérum	Non précisé	446
Gu F <i>et al.</i> , 2016	Tabac	Statut tabagisme N (%) : EAGLE: actuels = 80 (80) ; anciens = 0 ; jamais = 20 (20). PLCO : actuels = 45 (8,8) ; anciens = 241 (47,4) ; jamais = 223 (43,8). ATBC: actuels = 200 (100) ; anciens = 0 ; jamais = 0. SPAS : actuels = 44 (53,0) ; anciens et jamais = 39 (47,0)	Fumeur actuel, ancien fumeur ou jamais fumeur	Sang	Non précisé	514 (EAGLE) ; 426 (PLCO) ; 485 (ATBC) et 590 (SPAS)
Playdon MC <i>et al.</i> , 2017	Qualité de l'alimentation	Médiane [IIQ] : HEI-2010 Score = 54 [50-59] sur 100 ; aMED score = 3 [2-4] sur 8 ; HDI = 2 [1-3] sur 7 ; BSD score = 13 [10-16] sur 25	12 mois avant la visite	Sérum	1316	Non précisé
Hernández-Alonso P <i>et al.</i> , 2019	Consommation de vin rouge	Non-consommateurs de vin rouge : ≤ 1 verre/jour ; consommateurs de vin rouge : > 1 verre/jour	Au cours de l'année précédente	Plasma	Non précisé	386
Hernández-Alonso P <i>et al.</i> , 2020	Apport en protéines	Médiane [IIQ], % énergie/j : apport total en protéines = 16,29 [14,52-18,25] ; apport en protéines animales = 10,84 [9,16-12,87] ; apport en protéines végétales = 5,29 [4,7-6,05]	Non précisé	Plasma	Non précisé	385

Tableau supplémentaire III.C.3.3. Description des expositions et des mesures métabolomiques (suite)

Auteur, année	Exposition			Métabolomique		
	types d'exposition	quantification	durée d'exposition	fluide	nombre de métabolites mesurés	nombre de métabolites annotés
Papandreou C <i>et al.</i> , 2019	Consommation de café	Médiane (étendue), mL/jour : 50 (0; 475)	Au cours des 12 derniers mois	Plasma	387	Non précisé
Drouin-Chartier JP <i>et al.</i> , 2021	Consommation de produits laitiers	Moyenne (ET) consommation de produits laitiers (g/j) : Produits laitiers totaux : PREDIMED à l'inclusion = 375 (221) ; PREDIMED à 1 an = 366 (214) ; Cohortes regroupées américaines = 301 (227). Produits laitiers à teneur réduite en matières grasses : PREDIMED à l'inclusion = 265 (223) ; PREDIMED à 1 an = 278 (217) ; Cohortes regroupées américaines = 249 (221). Produits laitiers concentrés entiers : PREDIMED à l'inclusion = 31 (26) ; PREDIMED à 1 an = 28 (23) ; Cohortes regroupées américaines = 27 (22). Produits laitiers fermentés : PREDIMED à l'inclusion = 111 (94) ; PREDIMED à 1 an = 106 (87) ; Cohortes regroupées américaines = 52 (53). Lait : PREDIMED à l'inclusion = 252 (185) ; PREDIMED à 1 an = 250 (185) ; Cohortes regroupées américaines = 240 (216). Yaourt : PREDIMED à l'inclusion = 81 (86) ; PREDIMED à 1 an = 78 (83) ; Cohortes regroupées américaines = 28 (47). Fromage : PREDIMED à l'inclusion = 30 (25) ; PREDIMED à 1 an = 28 (23) ; Cohortes regroupées américaines = 24 (21)	PREDIMED : chaque année ; cohortes américaines : tous les 2 à 4 ans	Plasma	Non précisé	385
Wang F <i>et al.</i> , 2022	Régime alimentaire à base de plantes	Moyenne (ET), Plant-based Diet Index (PDI) = 55,2 (5,9) ; Healthful Plant-based Diet Index (hPDI) = 55,6 (6,8) ; Unhealthy Plant-based Diet Index (uPDI) = 54,3 (6,8)	Non précisé	Plasma	Non précisé	264
Tessier AJ <i>et al.</i> , 2024	Mode de vie sain	Score, Moyenne (ET) : 1,9 (1,1)	Sur 28 ans	Plasma	Non précisé	243
Huang T <i>et al.</i> , 2022	Abus précoces	Echantillon 1 : aucun = 263 (32,8%) ; abus léger = 91 (11,3%) ; abus modéré = 294 (36,6%) ; abus sévère = 155 (19,3%). Echantillon 2 : aucun = 55 (26,1%) ; abus léger = 12 (5,7%) ; abus modéré = 81 (38,4%) ; abus sévère = 63 (29,9%)	Enfance, adolescence (avant l'âge de 18 ans)	Plasma	283	Non précisé
Tabung FK <i>et al.</i> , 2020	Régime alimentaire	EDIP score, Echantillon de découverte : quantile 1 (-4,99; < -0,76) ; quantile 2 (60,76; < -0,18) ; quantile 3 (-0,18; < 0,28) ; quantile 4 (0,28; < 0,77) ; quantile 5 (0,77; 5,14). Echantillon de réplication : quantile 1 (-4,23; < -0,75) ; quantile 2 (-0,75; < 0,21) ; quantile 3 (-0,21; < 0,30) ; quantile 4 (0,30; 0,74) ; quantile 5 (0,74; 3,63)	Au cours des 3 derniers mois	Plasma	509	448
Che X <i>et al.</i> , 2024	Adhésion au régime de type Méditerranéen	Score MSDS : premier tertile (score < 24) : N = 468 ; deuxième tertile (score 24-27) : N = 450 ; troisième tertile (≥ 27) : N = 492	Pendant la grossesse	Plasma	Non précisé	432

Tableau supplémentaire III.C.3.3. Description des expositions et des mesures métabolomiques (suite)

Auteur, année	Exposition		durée d'exposition	fluide	nombre de métabolites mesurés	nombre de métabolites annotés
	types d'exposition	quantification				
Choi JJ <i>et al.</i> , 2023	Tabac	Statut tabagisme, N (%) : jamais: 711 (56,2) ; anciens: 469 (37,0) ; actuels: 86 (6,8)	Jamais-fumeurs, anciens fumeurs ou fumeurs actuels	Liquide céphalorachidien et plasma	1138 (plasma) and 390 (liquide céphalorachidien)	Non précisé
Chen X <i>et al.</i> , 2022	Pollution de l'air	Concentrations moyennes quotidiennes en période préconceptionnelle, médiane [IIQ] ($\mu\text{g}/\text{m}^3$) : $\text{PM}_{2,5}$: Contrôles = 47,8 [40,9-56,5] ; Cas = 46,7 [39,8-57,1]. PM_{10} : Contrôles = 80,1 [65,8-91,1] ; Cas = 76,6 [63,3-91,1]. SO_2 : Contrôles = 12,1 [10,8-14,5] ; Cas = 12,5 (11,1-14,8). NO_2 : Contrôles = 47,8 [33,3-61,9] ; Cas = 42,5 [30,0-60,1]. CO : Contrôles = 0,9 [0,8-1,0] ; Cas = 0,9 [0,8-1,1]. O_3 : Contrôles = 39,4 [26,5-53,8] ; Cas = 39,1 [24,2-54,6]	De la période préconceptionnelle jusqu'au troisième trimestre (estimé à partir des données 2015-2017)	Cheveux	Non précisé	276
Salihovic S <i>et al.</i> , 2019	PFAS	Médiane [IIQ] ng/mL : PFHxS = 2,08 [1,61-3,45] ; PFOS = 13,35 [10,13-17,79] ; PFHpA = 0,07 [0,05-0,11] ; PFOA = 3,33 [2,55-4,39] ; PFNA = 0,71 [0,53-0,97] ; PFUnDA = 0,29 [0,26-0,40]	Non clair (mesure réalisée sur des échantillons sanguins, prélevés le matin après un jeûne nocturne)	Sang	7522	203
Cohen CC <i>et al.</i> , 2022	Consommations de boissons sucrées	Quartile 1 (0 à 0,25 portions/j) : Moyenne (ET) (portions/j) = 0,11 (0,09) ; Quartile 2 (0,26 à 0,54 portions/j) : Moyenne (ET) (portions/j) = 0,39 (0,08) ; Quartile 3 (0,55 à 1,00 portions/j) : Moyenne (ET) (portions/j) = 0,77 (0,13) ; Quartile 4 (1,01 à 5,12 portions/j) : Moyenne (ET) (portions/j) = 1,86 (0,91)	Durant la jeunesse	Plasma	767	Non précisé
Alcolea JA <i>et al.</i> , 2023	Trihalométhanes et nitrates dans l'eau potable	Moyenne (ET) : THMs totaux ($\mu\text{g}/\text{L}$): contrôles = 37,1 (25,0) ; cas = 42,9 (33,6). Brominated (Br) THMs ($\mu\text{g}/\text{L}$) : contrôles = 17,2 (21,8) ; cas = 27,5 (30,6). Chloroform ($\mu\text{g}/\text{L}$) : contrôles = 19,9 (13,1) ; cas = 15,4 (11,0). Nitrate (mg/L) : contrôles = 5,4 (5,5) ; cas = 7,1 (7,4)	Moyenne annuelle	Sérum	1629	26

Tableau supplémentaire III.C.3.3. Description des expositions et des mesures métaboliques (fin)

Auteur, année	Exposition			Métabolomique		
	types d'exposition	quantification	durée d'exposition	fluide	nombre de métabolites mesurés	nombre de métabolites annotés
Walker ME <i>et al.</i> , 2020	Profils alimentaires sains	Moyenne \pm ET. AHEI (0-110) : hommes = 48,0 \pm 9.6 ; femmes = 51,7 \pm 9,7. DASH (0-40) : hommes = 23,0 \pm 4,9 ; femmes = 25,3 \pm 4,7. MDS (0-25) : hommes = 12,3 \pm 4.3 ; femmes = 12,0 \pm 4,4	Durant l'année qui a précédé la visite de la Heart Study	Plasma	216	Non précisé
Zheng Y <i>et al.</i> , 2014 (a)	Habitudes de consommation d'alcool	N (%). Buveurs occasionnels : échantillon de découverte = 110 (7,3) ; échantillon de réplication = 34 (7,1). Petits buveurs : échantillon de découverte = 110 (7,3) ; échantillon de réplication = 37 (7,8). Buveurs faibles à modérés : échantillon de découverte = 105 (7,0) ; échantillon de réplication = 39 (8,2). Buveurs modérés à élevés : échantillon de découverte = 121 (8,1) ; échantillon de réplication = 25 (5,2)	Lors de l'examen initial (fréquence habituelle de consommation de vin, de bière ou de spiritueux)	Sérum	Non précisé	356
Zheng Y <i>et al.</i> , 2014 (b)	Habitudes alimentaires	Apport énergétique total, Moyenne (ET), calories par jour : échantillon de découverte = 1 574,00 (639,4) ; échantillon de réplication = 1 579,50 (568,4)	Non précisé	Sérum	Non précisé	356
Beuchel C <i>et al.</i> , 2019	Facteurs cliniques et liés au mode de vie	Statut tabagique (actuel/ancien/jamais), N (%) : LIFE-adult = 2034 (21,5)/2706 (28,5)/4483 (47,3) ; LIFE-Heart = 1581 (27,4)/2108 (36,6)/2063 (35,8) ; Sorbs = 150 (15,4)/195 (20)/616 (63,2)	Jamais-fumeurs, anciens fumeurs ou fumeurs actuels	Sang	63	Non précisé
Wang Y <i>et al.</i> , 2020	Lieu de résidence urbain ou rural	Non précisé	Non précisé	Plasma	1108	725 annotés au niveau 1 ; 165 annotés au niveau 2 ; 17 annotés au niveau 3 ; 200 annotés au niveau 4
Zhang R <i>et al.</i> , 2022	Tabac	Statut tabagique (N) : actuel = 248 ; ancien = 367 ; jamais = 637	Jamais-fumeurs, anciens fumeurs ou fumeurs actuels	Sérum	1202	Non précisé

Zheng Y *et al.*, 2014 (a) : Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study

Zheng Y *et al.*, 2014 (b) : Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study

Abréviations : IIQ = intervalle interquartile ; ET = écart type ; PM_{2,5} = particules en suspension ayant un diamètre aérodynamique inférieur à 2,5 μ m ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 μ m ; SO₂ = dioxyde de soufre ; CO = monoxyde de carbone ; O₃ = ozone ; PFHxS = perfluorohexane sulfonic acid ; PFOS = perfluorooctane sulfonate ; PFHpA = perfluoroheptanoic acid ; PFOA = perfluorooctanoic acid ; PFNA = perfluorononanoic acid ; PFUnDA = perfluoroundecanoic acid ; THMs = trihalométhanes

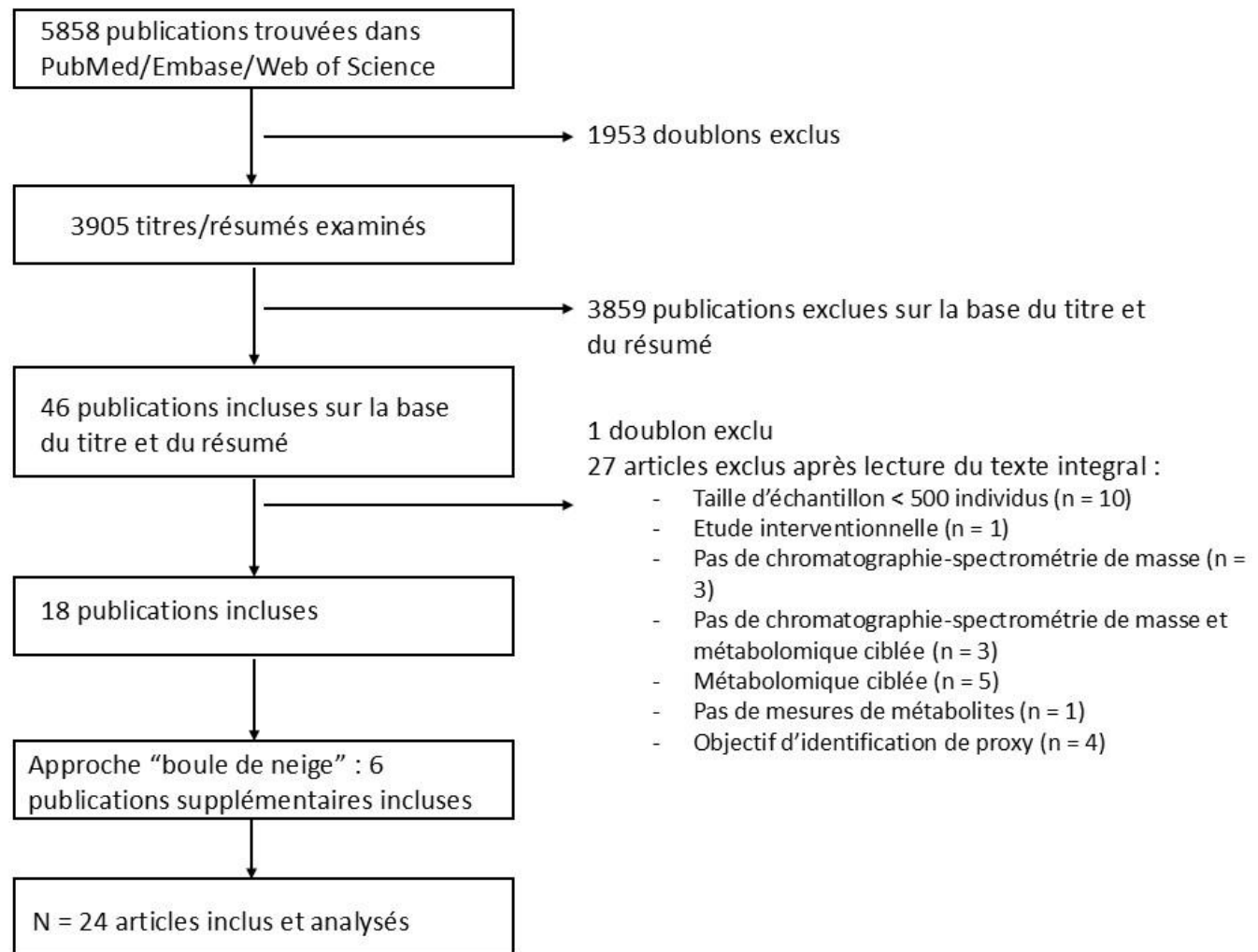


Figure III.C.3.1 : Diagramme de flux (Flow chart)

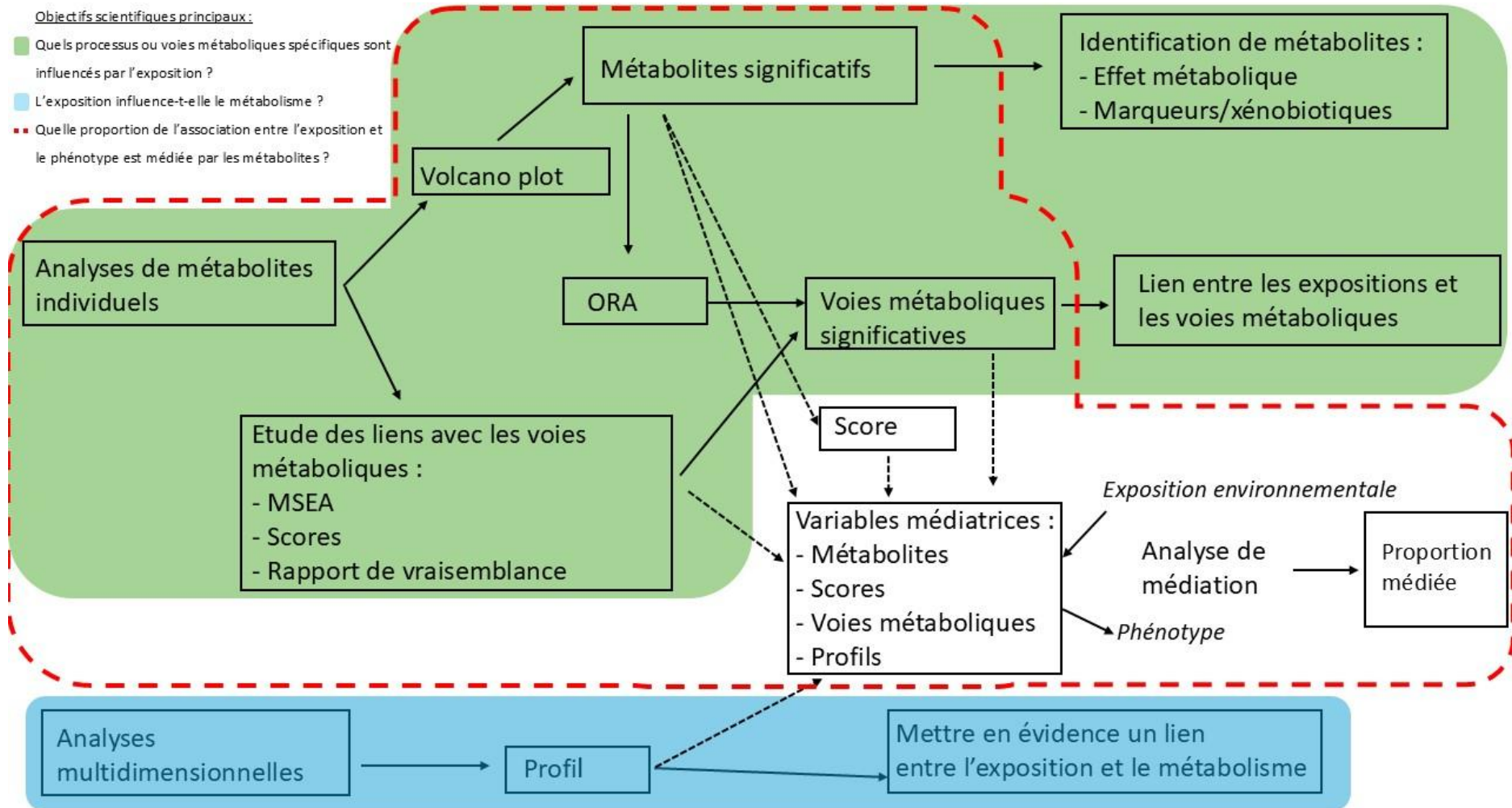


Figure III.C.3.2 : Cadre proposé pour adapter l'analyse statistique à l'objectif de l'étude

4. Discussion de cette étude

a) Résultats principaux

Dans cette revue méta-épidémiologique des associations entre expositions environnementales et métabolomique, 24 articles ont satisfait nos critères d'inclusion. Au total, 2 études (221,231) remplissaient nos trois critères : une hypothèse H1 pertinente au regard de la question scientifique, un outcome statistique pertinent au regard de l'hypothèse H1, et l'absence de problème statistique identifié. Il convient de noter que le score ROBINS-I n'inclut pas d'items relatifs à la validité des tests statistiques ni à la définition des objectifs d'étude, et n'a donc pas permis d'identifier les problèmes méthodologiques complexes présents dans les 22 autres études.

b) Représentations graphiques

Les représentations graphiques retrouvées dans les études sont hétérogènes. Alors que certaines sont fréquemment utilisées (forest plot, heatmap, barplot, volcano plot), beaucoup d'autres apparaissent plus sporadiquement (par exemple, diagramme en barres horizontales, Manhattan plot ou bubble plot). Il pourrait être intéressant d'établir un consensus sur les représentations graphiques des associations entre expositions environnementales et métabolites ou voies métaboliques. L'intérêt d'avoir un consensus est de pouvoir comparer facilement et rapidement les différentes études. De plus, cela permettrait d'éviter que les représentations graphiques soient sélectionnées a posteriori en fonction des résultats.

Par ailleurs, certaines représentations (par exemple, diagramme de Venn (230), forest plot (215)) ne montrent que les métabolites significativement associés. Il pourrait être pertinent de disposer également d'une représentation de l'ensemble des métabolites (associés et non associés) afin d'avoir une meilleure idée du nombre d'associations significatives par rapport au nombre total de métabolites mesurés.

C'est le cas, par exemple, du volcano plot, qui permet de visualiser à la fois tous les métabolites et ceux situés au-dessus du seuil de significativité. Dans le volcano plot, l'axe des abscisses représente les coefficients d'association entre chaque métabolite et l'exposition testée, tandis que l'axe des ordonnées représente la valeur $-\log_{10}$ de p . Chaque métabolite est représenté par un point, et tous les métabolites testés figurent sur le graphique. Le seuil de significativité peut être indiqué par une ligne horizontale, ce qui facilite l'identification des métabolites significativement associés à l'exposition (ceux situés au-dessus de ce seuil).

En outre, d'autres études non incluses dans cette revue utilisent parfois le diagramme quantile-quantile ou QQ plot, qui permet également de visualiser l'ensemble des métabolites testés (235). Il représente la distribution attendue des p (en abscisse) et la distribution observée (en ordonnée). Une valeur λ est associée au QQ plot et exprime l'écart entre la distribution observée et la distribution attendue du test (la diagonale). Par exemple, une valeur de λ égale à 1 traduit une distribution observée similaire à celle attendue. Si un grand nombre de valeurs de p se trouvent à des niveaux plus faibles que prévu, cela suggère la présence d'associations (même si aucune de ces associations n'est significative après correction pour tests multiples). L'intérêt de cette approche est de visualiser l'écart des p issus de l'ensemble des tests effectués sur les métabolites par rapport à ce qui serait attendu en

l'absence d'effet de l'exposition, et ainsi de détecter des associations. Une limite est qu'un QQ plot fortement dévié peut aussi s'expliquer par de fortes corrélations entre variables (« structure de population »), sans que les associations soient significatives. Actuellement, aucun test statistique n'est associé au QQ plot pour trancher entre ces deux hypothèses.

Nous proposons d'utiliser un test du rapport de vraisemblance qui, s'il est significatif, indiquerait une différence significative liée à l'exposition, et à l'inverse, s'il ne l'est pas, suggérerait une « structure de population ». Le test du rapport de vraisemblance implique un grand nombre de degrés de liberté, ce qui le rend plus adapté lorsque le nombre de variables n'est pas trop élevé. Ainsi il semble plus adapté aux représentations graphiques concernant les voies métaboliques que les métabolites individuels.

c) La relation entre la question scientifique et le type d'analyse

(1) Question scientifique : « Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? » - analyses unidimensionnelles.

Une méthodologie statistique appropriée pour ce type de question scientifique devrait inclure une analyse au niveau des voies métaboliques (par exemple, une méthode d'enrichissement) éventuellement complétée d'une analyse des métabolites individuels (par exemple, des régressions linéaires avec correction pour le taux de fausses découvertes (FDR)).

Analyses de l'ensemble des métabolites ou uniquement des métabolites annotés

Quatorze des études incluses ont analysé uniquement les métabolites annotés (64,65,212,214,216–218,220,222–224,226–228), tandis que les dix autres ont analysé l'ensemble des métabolites mesurés (66,213,215,219,221,225,229–232). L'analyse de l'ensemble des métabolites permet d'exploiter toutes les informations disponibles. Cependant, lorsqu'un métabolite non identifié est retrouvé significativement associé à l'exposition, son identification peut s'avérer difficile, voire impossible. Il devient alors complexe de tirer des conclusions sur les conséquences physiopathologiques de l'exposition (80). Par ailleurs, l'analyse de l'ensemble des métabolites nécessite des corrections pour les tests multiples très strictes, ce qui abaisse le seuil de significativité et accroît le risque de faux négatifs. À l'inverse, l'analyse limitée aux seuls métabolites annotés permet d'appliquer des corrections moins stringentes, d'augmenter la puissance statistique et donc d'accroître la probabilité d'identifier des métabolites d'importance mécanistique.

Il reste toutefois possible de recourir à des approches analytiques complémentaires afin d'identifier des métabolites associés mais non annotés.

Analyses des métabolites individuels

Les modèles statistiques utilisés dans les analyses de métabolites dépendaient de la nature de l'exposition environnementale (qualitative ou quantitative). Cette approche permet d'identifier les métabolites associés à l'exposition, mais elle implique un grand nombre de tests statistiques et donc des corrections pour tests multiples – ce qui accroît la probabilité de faux

négatifs. Dans la mesure où seuls les métabolites sont identifiés, la détermination des mécanismes sous-jacents reste difficile.

Une analyse des métabolites individuels peut précéder une analyse de voies métaboliques. Toutefois, le recours à un nombre important de tests peut réduire la puissance de l'étude et conduire à l'absence d'association détectée avec les métabolites.

Analyses basées sur les voies métaboliques

Sept des études portant sur les voies métaboliques ont appliqué une méthode d'enrichissement (65,66,215,228,230–232). Les deux méthodes d'enrichissement les plus utilisées étaient l'ORA (*over-representation analysis*, dans cinq études) (66,228,230–232) et la MSEA (dans deux études) (65,215).

Une étude a appliqué l'ORA uniquement aux protéines, et non aux métabolites (230). Les outils bio-informatiques utilisés dans ces études étaient variés : MetaboAnalyst (66,228,231), mummichog (66), WebGestalt (230), et Metabolon (232). L'ORA permet de déterminer si les métabolites significativement associés à une exposition sont sur-représentés (236). Par exemple, l'entrée de MetaboAnalyst correspond à la liste des métabolites significativement associés à l'exposition. L'algorithme cartographie ensuite les voies métaboliques impliquées et analyse la sur-représentation ainsi que la topologie des voies (79). L'une des principales limites de l'ORA est qu'elle nécessite que les métabolites soient significativement associés à l'exposition après correction pour tests multiples. En effet, la liste des métabolites doit être filtrée en fonction de la valeur de p et du *fold change* (le *fold change* correspondant à l'exponentielle du coefficient d'association entre l'exposition et le métabolite). Cette approche est confrontée à des défis liés à une faible puissance statistique et à une forte probabilité d'associations faussement positives ou faussement négatives. De plus, des effets de seuil peuvent affecter les métabolites dont la valeur de p est proche du seuil de significativité, et leur inclusion ou exclusion de l'analyse d'enrichissement peut être fortement influencée par des facteurs aléatoires.

La MSEA (*metabolite set enrichment analysis*) prend en compte l'ensemble des métabolites (qu'ils soient ou non significativement associés à l'exposition), classe les associations et détermine ainsi quelles voies métaboliques sont sur-représentées. Deux études ont appliqué cette méthode (65,215). La MSEA contourne les limites liées aux critères de sélection (valeur de p et *fold change*) en établissant une liste ordonnée des métabolites (par exemple, *signe du fold change* $\times -\log_{10}(p\text{-value})$). Cette approche intègre l'ensemble des coefficients d'association, les hiérarchise et identifie les voies métaboliques dans lesquelles les métabolites les plus fortement associés à l'exposition sont sur-représentés (237).

En théorie, la MSEA est limitée par sa dépendance aux comparaisons de rangs dans le test d'association. Par définition, toutes les voies métaboliques ne peuvent pas être sur-représentées dans les rangs extrêmes. Pour une exposition ayant un impact étendu sur de multiples voies métaboliques (par exemple l'obésité), des associations réelles avec certaines voies peuvent être masquées par celles présentant des associations plus fortes.

Dans les deux méthodes d'enrichissement, les tests sont limités aux voies pouvant être explorées à partir des métabolites annotés : les métabolites non annotés ne peuvent pas être inclus. Cela réduit considérablement le nombre de tests statistiques réalisés et minimise l'impact des corrections pour comparaisons multiples. Il convient de noter que, bien que

certaines outils bio-informatiques permettent de réaliser une ORA sur l'ensemble des métabolites, l'inclusion de métabolites non annotés rend ces analyses incomplètes.

Les méthodes d'analyse de voies métaboliques sont dérivées de celles utilisées en génétique. Or, les données génétiques comprennent actuellement un nombre bien plus important de variables et de gènes par voie que ce n'est le cas pour les métabolites. Il reste donc à déterminer si ces approches sont réellement optimales pour la métabolomique. Le nombre plus restreint de variables (et donc de degrés de liberté) dans les analyses de voies métaboliques par rapport aux voies génétiques pourrait permettre de développer et d'utiliser des modèles statistiques plus simples.

Problèmes statistiques

Notre revue a identifié une autre méthode d'analyse de voies : la méthode de Fisher. Les deux articles ayant utilisé cette approche (225,232) n'ont pas précisé comment la méthode avait pu être adaptée. Or, la méthode de Fisher repose sur l'hypothèse d'indépendance des tests statistiques combinés, ce qui n'est pas adapté dans les jeux de données métabolomiques où les métabolites sont fortement corrélés entre eux. Par ailleurs, une étude n'a pas spécifié la nature de la méthode d'analyse de voies employée (226).

Autres tests possibles

Wieder *et al.*, ont appliqué une analyse de voies à échantillon unique (*single sample pathway analysis*, ssPA), qui attribue un score spécifique à chaque voie métabolique au niveau individuel (238). La ssPA repose sur les connaissances préalables concernant l'appartenance des métabolites à une liste de référence de voies métaboliques. Un score est alors attribué à chaque voie et résume son niveau d'activation. Les trois principaux types de ssPA sont : (i) les méthodes de réduction de dimension (telles que la décomposition en valeurs singulières, qui extrait des composantes à partir du jeu de données), (ii) les méthodes basées sur la GSEA (comme la GSEA à échantillon unique (ssGSEA)), et (iii) les méthodes basées sur un score qui normalise les données (par exemple le Z-score) (238). L'avantage de la ssPA est qu'elle réduit le nombre de tests statistiques effectués (le nombre de tests correspond au nombre de voies métaboliques) et, par conséquent, augmente le seuil de significativité. Un autre avantage de la ssPA (comme pour la MSEA) est la possibilité de déterminer si les métabolites d'une voie sont collectivement associés à une exposition, même lorsque chacun des métabolites pris isolément n'a pas d'effet significatif.

L'inconvénient de la décomposition en valeurs singulières réside dans la difficulté d'interprétation des composantes principales extraites. Le principal inconvénient de l'approche par Z-score est qu'elle indique seulement si l'exposition est associée à une augmentation ou une diminution des niveaux de métabolites dans une voie métabolique ; elle ne permet pas d'identifier un mécanisme dont le blocage pourrait entraîner une augmentation de certains métabolites de la voie et une diminution d'autres. La méthode de Fisher peut être adaptée pour tenir compte de la non-indépendance des tests, comme le suggèrent Dai *et al.*, (234).

Nous suggérons que le test du rapport de vraisemblance (*TRV*) pourrait également présenter un intérêt. Ce test compare deux modèles dans lesquels l'exposition environnementale est la variable explicative : l'un inclut uniquement les variables d'ajustement et l'autre inclut, en plus, les niveaux des métabolites d'une voie donnée. Le test du rapport de vraisemblance permet d'évaluer si l'ajout, des métabolites appartenant à une voie, dans le modèle permet de mieux prédire l'exposition de façon statistiquement significative. Cette approche permettrait de tester une association potentielle (présumée) entre les voies métaboliques et une exposition environnementale, sans faire d'hypothèse sur la direction de l'association de chaque métabolite individuel au sein de la voie.

Le TRV implique un nombre élevé de degrés de liberté (chaque métabolite ajoutant un degré de liberté au modèle), ce qui le rend plus adapté lorsque le nombre de variables n'est pas trop important. Le rapport de vraisemblance pourrait être particulièrement pertinent dans les analyses où chaque voie est résumée par un score, maintenant ainsi le nombre de variables à un niveau gérable. Dans ce cas, l'exposition est la variable dépendante dans les deux modèles statistiques. Le premier modèle inclurait uniquement les variables d'ajustement, tandis que le second modèle comprendrait en plus les scores des voies. Le TRV permettrait alors d'évaluer si le modèle intégrant les voies explique mieux l'exposition que le modèle sans celles-ci.

Les méthodes d'analyse métabolomique sont dérivées de celles utilisées en génétique. Toutefois, le très grand nombre de gènes implique également un nombre très élevé de degrés de liberté, ce qui peut expliquer pourquoi le TRV n'a pas encore été largement appliqué. En métabolomique, le nombre relativement limité de degrés de liberté permet d'envisager l'utilisation de méthodes plus simples, comme le TRV.

Corrections des tests multiples

Pour les métabolites individuels, toutes les études examinées ont corrigé leurs analyses pour les tests multiples en utilisant soit la méthode de Bonferroni, soit le FDR (*false discovery rate*). Les analyses de voies métaboliques ont principalement eu recours au FDR, car il est automatiquement intégré dans les outils bio-informatiques d'analyse d'enrichissement de voies (239). Cette méthode contrôle le FDR en ajustant les valeurs de significativité (*p-values*) (240). L'avantage du FDR est qu'il est moins stringent que la méthode de Bonferroni. Néanmoins, son application reste associée à une réduction marquée de la puissance statistique.

Autres tests possibles

Les méthodes utilisées pour corriger les tests multiples sont bien établies et fiables. Cependant, d'autres approches pourraient être envisagées, comme celle décrite par Storey *et al.* Au lieu de fixer α au préalable puis d'estimer le nombre de rejets, Storey *et al.*, proposent de définir d'abord la région de rejet puis d'estimer α . Cette méthode pourrait présenter une puissance statistique supérieure à celle des méthodes conventionnelles (jusqu'à dix fois plus élevée que le FDR, par exemple) (241,242).

Synthèse sur cet objectif

Pour les 13 études visant à comprendre les mécanismes biologiques affectés par l'exposition environnementale, 10 ont examiné les métabolites individuellement. Cela a permis d'obtenir une première compréhension des mécanismes. Toutefois, au regard de cette question scientifique, une analyse au niveau des voies métaboliques serait plus pertinente et plus complète. Néanmoins, seules 4 des 13 études ont appliqué une analyse de voies métaboliques (225,228,230,231). Trois études ont réalisé une ORA (228,230,231), dont une portant uniquement sur des protéines (et non sur des métabolites) (230). Une étude a utilisé un test de Fisher classique (225), dont la validité peut être remise en question car les tests statistiques n'étaient pas nécessairement indépendants. D'autres méthodes seraient plus appropriées, telles que la MSEA, un test de Fisher modifié, une approche basée sur des scores, ou une approche basée sur un TRV.

(2) Question scientifique : « Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ? » – calcul de la proportion médiée

Pour cet objectif, les méthodes statistiques appropriées devraient inclure le calcul de la proportion médiée. En réalité, seules cinq des dix études (65,215,219,221,223) ont calculé la proportion de l'association entre l'exposition et le phénotype qui était médiée par les métabolites. Parmi ces cinq études, trois ont utilisé un score pondéré (65,215,223), une a sélectionné les métabolites associés à l'exposition et les a inclus dans l'analyse de médiation (219), et la dernière a utilisé les composantes principales des métabolites (221) comme mesures de la médiation.

Problèmes statistiques

Dans les quatre études ayant utilisé un score pondéré ou les métabolites associés à l'exposition (65,215,219,223), l'échantillon ayant servi à construire le score ou à sélectionner les variables a également été inclus dans l'analyse de médiation.

Lorsque l'on teste l'effet médiateur, l'hypothèse nulle correspond soit à l'absence d'association entre l'exposition et le médiateur, soit à l'absence d'association entre le médiateur et le critère de jugement (243). Dans les quatre études, le médiateur était défini soit (i) par un score construit à partir de l'association entre l'exposition et les métabolites, avec un calcul de la médiation effectué dans le même échantillon que celui utilisé pour construire le score, soit (ii) par une sélection des seuls métabolites significativement associés à l'exposition, l'analyse de médiation étant réalisée sur ce même échantillon. Dans les deux cas, cela peut conduire à un surapprentissage (*overfitting*). En d'autres termes, le modèle peut être trop étroitement ajusté à l'échantillon d'entraînement, et les métabolites (ou les pondérations définies pour le score) peuvent avoir été sélectionnés (au moins en partie) du fait de variations aléatoires propres à l'échantillon ; ceci pourrait conduire à une surestimation de l'effet médiateur des métabolites. Ainsi, la définition des variables de médiation ne devrait pas dépendre de l'exposition ou du phénotype, et l'utilisation d'échantillons distincts de construction et de validation peut être nécessaire. Ces problèmes sont évités par l'utilisation de l'ACP, puisque les composantes sont définies indépendamment de l'exposition et du phénotype.

Autres tests possibles

Des méthodologies alternatives développées pour l'analyse de médiateurs multiples peuvent être utilisées (244).

La modélisation par équations structurelles apparaît particulièrement pertinente pour les analyses de médiation (245). En effet, ces méthodes permettent d'évaluer les relations entre les expositions environnementales et les problèmes de santé de manière indirecte, en examinant la proportion médiée par les perturbations métaboliques. La modélisation par équations structurelles peut identifier les métabolites médiateurs ainsi que les voies métaboliques impliquées. La seule étude ayant eu recours à cette approche avait toutefois considéré l'indice de masse corporelle comme médiateur potentiel, et non les métabolites (228). Il serait intéressant d'utiliser la modélisation par équations structurelles pour explorer la proportion d'une relation entre exposition et santé médiée par les métabolites.

Synthèse sur cet objectif

Parmi les dix études ayant formulé un objectif correspondant à une médiation, cinq ont calculé un effet médié. Parmi ces cinq études, une seule ne présentait aucun problème statistique (par exemple de surapprentissage ou overfitting).

(3) Analyses multidimensionnelles – question scientifique :
« déterminer si l'exposition altère le métabolisme global ».

Certaines méthodologies statistiques ne sont pas mentionnées dans les paragraphes précédents, en particulier dans les sections consacrées aux mécanismes physiopathologiques et à la médiation, car elles n'étaient pas directement liées aux hypothèses H1 testées. C'est notamment le cas des analyses multidimensionnelles. Ces méthodologies statistiques peuvent contribuer à répondre à des objectifs liés à la compréhension des mécanismes biologiques et à la médiation, mais également à d'autres objectifs.

Douze des études examinées ici ont utilisé diverses méthodes analytiques en lien avec la forte dimensionnalité des données (65,66,212–214,220–223,228,229,232). Celles-ci incluaient des méthodes de régression pénalisée, telles que la régression LASSO et la régularisation elastic net. La régression LASSO permet de réduire utilement la dimensionnalité en sélectionnant des variables et en appliquant une pénalité aux coefficients, ce qui revient à en diminuer le poids, pouvant aller jusqu'à zéro. Cependant, la régression LASSO présente des limites : lorsque les variables sont fortement corrélées, le modèle en sélectionne une de manière arbitraire, et ce choix peut ne pas être optimal. De plus, le nombre de variables non pénalisées est limité par la taille de l'échantillon (246–248).

La régularisation elastic net combine de manière utile les pénalités issues à la fois de la régression LASSO et de la régression ridge.

La régression ridge gère efficacement les variables corrélées, mais ne permet pas de réduire la dimensionnalité, car la pénalité appliquée aux variables ne les réduit pas à zéro. Comme la régression LASSO, la régularisation elastic net ne permet pas de gérer de manière totalement fiable des données fortement corrélées (246–248). Ces méthodes ont été rarement

rapportées dans les publications examinées ici. En effet, la régression LASSO n'a été utilisée que dans une seule étude (229), tandis que la régularisation elastic net a été appliquée par un groupe issu du même consortium (incluant la cohorte espagnole PREDIMED et des cohortes américaines) (65,212–214,222,223).

La méthode LASSO permet d'identifier les métabolites associés à l'exposition. De même que les méthodes intégrant une correction pour tests multiples, ces approches (LASSO et elastic net) peuvent contribuer à identifier des mécanismes physiopathologiques.

Par ailleurs, des analyses de type ACP (212,221,232) et PLS (66,232) ont été utilisées. Ces méthodes visent à réduire la dimensionnalité des données. L'ACP génère des dimensions indépendamment de l'exposition, puis l'association de ces dimensions avec l'exposition est ensuite testée. Une limite de l'ACP est que les dimensions sont construites indépendamment de l'exposition et peuvent donc ne pas être les plus pertinentes pour l'étude de l'exposition. À l'inverse, la PLS prend en compte l'exposition, mais il n'existe pas de méthode simple permettant d'obtenir une valeur de p pour tester une hypothèse H_1 .

Enfin, certaines méthodes n'ont été utilisées que dans une seule étude : l'UMAP (220), la classification par forêts aléatoires (232), l'analyse d'intégration N (66), et la MMLR (228). Enfin, la méthode DIABLO, utilisée pour l'analyse d'intégration de type N-integration, permet l'étude simultanée de l'exposition et du phénotype au sein d'un même modèle (interactions multidimensionnelles). Cette méthode peut être utilisée pour explorer des mécanismes biologiques ou pour réaliser des analyses de médiation. Cependant, l'article qui l'a appliquée (66) n'a pas clairement extrait la composante médiée.

L'avantage des méthodes multidimensionnelles est qu'elles nécessitent une puissance statistique moindre. Leur principal inconvénient, pour certaines méthodes (par exemple la PLS) — à l'exception de la MMLR, qui permet de réaliser un test statistique — est l'absence de valeurs de p et une absence de contrôle du taux de faux positifs ; il existe ainsi un risque de faible reproductibilité et d'identification de profils métaboliques associés à l'exposition uniquement en raison des fluctuations d'échantillonnage. Pour évaluer la reproductibilité des résultats, il est possible d'appliquer une validation croisée ou d'utiliser un échantillon de validation en complément d'un échantillon d'apprentissage. C'était le cas dans neuf des douze études ayant appliqué des analyses multidimensionnelles (65,66,212–214,222,223,229,232).

Une autre limite concerne l'interprétation physiologique des résultats. Les dimensions générées ne tiennent pas compte des voies métaboliques dans leur construction, et l'interprétation physiologique des résultats peut ne pas être évidente. Cela peut ne pas poser de problème lorsque la question scientifique est simplement de savoir si l'exposition est associée au métabolisme, sans focus spécifique sur la physiopathologie. Par exemple, cette approche peut être pertinente pour déterminer si un xénobiotique a ou non un impact biologique. Dans ce type de situation, la réduction de dimension peut augmenter la puissance des études. Dans le cadre de notre méthodologie, nous n'avons identifié aucune étude correspondant à cette question scientifique ; néanmoins, ces méthodologies peuvent être utiles dans de tels cas.

Problèmes statistiques

Aucun problème statistique n'a été identifié dans ces méthodes.

Synthèse sur cet objectif

Des analyses multidimensionnelles ont été réalisées, mais il n'est pas toujours possible de les relier aux objectifs et aux hypothèses que nous avons identifiés. Elles peuvent être utiles pour l'étude de la médiation et des aspects physiopathologiques, mais elles présentent également un avantage supplémentaire : en réduisant le nombre de dimensions, elles peuvent contribuer à répondre à la question suivante : « L'exposition influence-t-elle le métabolisme ? »

d) Forces et limites de cette revue

L'évaluation de la cohérence entre l'objectif de l'étude et la méthodologie statistique, ainsi que la validité des tests statistiques, nous a permis d'identifier et de classer des écueils méthodologiques qui ne sont pas pris en compte par des outils d'évaluation du biais tels que le score ROBINS-I. Toutefois, les critères relatifs à la question scientifique de l'étude, à l'adéquation de la méthode statistique à cet objectif et à l'identification des problèmes statistiques n'étaient pas inclus dans le protocole initial de cette revue. En effet, l'hétérogénéité des questions scientifiques des études et des analyses statistiques a rendu très difficile la synthèse des résultats, ce qui nous a conduits à définir un critère de cohérence méthodologique entre les études.

Une autre limite est que, bien que ces indicateurs reflètent des exigences fondamentales du design des études, leur interprétation peut dépendre de la subjectivité des auteurs de la revue. Nous n'avons pas identifié, avant de débiter cette revue systématique, de cadre adapté permettant de classer et d'évaluer la méthodologie de ces études. Le score ROBINS s'est avéré peu pertinent dans ce contexte. Les critères d'inclusion et l'extraction des données n'ont pas été modifiés par rapport à la méthodologie initiale. Toutefois, nous avons dû construire a posteriori un cadre plus pertinent pour interpréter les données collectées, incluant une évaluation de la cohérence entre la question scientifique, l'hypothèse H1 et l'analyse statistique.

Nous n'avons pas été en mesure de définir a priori la classification des questions scientifiques, ainsi que les règles d'évaluation de la cohérence. Comparée à des scores de qualité plus génériques, cette approche peut permettre de formuler des conclusions plus précises et de mieux orienter les améliorations des études futures. Toutefois, la méthodologie a pu être en partie guidée par les données, ce qui peut avoir affaibli la robustesse des résultats.

Pour évaluer la cohérence entre la question scientifique et la méthodologie, nous avons limité l'analyse à des sections spécifiques de l'article (par exemple, la fin de l'introduction pour l'hypothèse H1). La force de cette approche est qu'elle évite de déduire les questions scientifiques à partir de la méthodologie et donc de conclure de manière artificielle à une cohérence entre les deux.

En revanche, cette approche a pu conduire à classer de manière trop sévère certains articles comme incohérents lorsqu'ils ne suivaient pas strictement les conventions d'écriture scientifique et qu'ils présentaient des informations pertinentes dans des sections jugées moins appropriées (par exemple, la discussion).

5. Conclusion

La présente revue a conduit à suggérer l'utilisation de différentes approches méthodologiques en fonction de l'hypothèse testée (Figure III.C.3.2). Des recommandations détaillées pour ces analyses restent toutefois à établir.

Lorsque l'objectif est de comprendre les mécanismes biologiques sous-jacents, il serait approprié de réaliser une analyse des métabolites individuels (avec correction pour tests multiples via le FDR ou la méthode de Storey), puis une analyse de voies métaboliques basée sur l'enrichissement (par exemple ORA ou MSEA). La MSEA est particulièrement utile car elle évite la nécessité de sélectionner au préalable des métabolites significatifs. D'autres méthodes (telles que les approches basées sur des scores ou sur des rapports de vraisemblance) peuvent également être pertinentes.

Lorsque l'objectif est d'étudier l'effet médiateur des métabolites dans l'association entre l'exposition et le phénotype, il est important d'utiliser une méthode permettant de quantifier la proportion médiée. Le métabolome peut être défini de plusieurs manières, et il est donc essentiel de s'assurer que la définition retenue est indépendante du modèle de médiation final.

Lorsque l'objectif est de déterminer si l'exposition a un impact sur le métabolisme (sans nécessairement chercher à comprendre les mécanismes biologiques — par exemple, dans le cas de l'étude d'un effet d'une exposition à faible dose à un polluant), une analyse multidimensionnelle (par exemple ACP ou PLS) permet d'éviter les tests multiples et de ne pas avoir à prendre en compte l'ensemble du métabolome. L'inconvénient est que l'étude du métabolisme dans son ensemble ne permet pas une compréhension directe des mécanismes impliqués. Les analyses multidimensionnelles utilisées dans la littérature sont hétérogènes, et notre revue n'a pas permis de hiérarchiser les différentes méthodes en termes d'utilité potentielle.

En conclusion, les résultats de la présente revue indiquent que (i) il est essentiel de définir clairement la question scientifique et l'hypothèse H1, (ii) il est essentiel de définir clairement des termes tels que « profils » et « signatures » lorsqu'ils sont utilisés, (iii) il est essentiel que la question scientifique, l'hypothèse H1 et l'outcome statistique soient cohérents, et (iv) la validité des tests statistiques est souvent compromise, notamment en raison de l'overfitting ; une attention particulière devrait être portée à ce point.

L'analyse métabolomique est un outil prometteur et puissant pour étudier les impacts de l'environnement sur la santé. Notre revue met en évidence une grande hétérogénéité des approches méthodologiques. La définition de la question scientifique et le choix de la méthodologie appropriée correspondante ne sont pas toujours simples, et des problèmes statistiques sont fréquents en raison de la complexité des analyses. Nous proposons un cadre méthodologique préliminaire pour l'évaluation et la conception de ces analyses, qui pourrait constituer une première étape vers l'élaboration d'un consensus ou de recommandations. Cela serait utile pour améliorer et standardiser les analyses métabolomiques et permettre aux chercheurs de tirer pleinement parti des informations générées par les techniques modernes de métabolomique.

D. Associations entre expositions environnementales et métabolomique

1. Introduction

Les expositions environnementales ont des impacts sur la santé. Cependant, des expositions à des doses faibles ou modérées peuvent avoir des effets sans conséquences cliniques identifiables à court terme mais, entraînant des impacts à long terme sur la santé. Ces effets sont difficiles à détecter et l'approche métabolomique non-ciblée est pertinente car elle n'est pas limitée par les hypothèses que l'on fait a priori et offre la possibilité de découvrir de nouveaux signaux (voir l'introduction générale pour plus de précisions).

La revue systématique a permis de proposer des stratégies d'analyses statistiques pour des études de métabolomique portant sur des échantillons de grande taille.

L'objectif était d'étudier les associations entre les expositions environnementales et les modifications des processus métaboliques, en population générale majoritairement faiblement exposée, afin de mieux comprendre les mécanismes biologiques sous-jacents aux faibles doses d'exposition ainsi que leurs effets potentiellement subcliniques.

Pour cela, nous allons étudier les expositions environnementales disponibles dans la base de données ELISABET, ainsi que les données de métabolomique disponibles par le projet Metalisabet.

2. Méthodes

Nous avons appliqué une transformation logarithmique à chaque métabolite, afin de s'approcher d'une distribution normale.

La concentration des métabolites dans les échantillons d'urine est affectée par la concentration des urines. Nous avons donc réalisé une analyse de littérature concernant les techniques de normalisation des niveaux des métabolites, pour tenir compte au mieux de la variabilité interindividuelle de la dilution des urines. Les études ont montré que normaliser avec le volume urinaire ou la créatinine urinaire ne permettait pas de corriger la variabilité interindividuelle. En revanche, les méthodes suivantes étaient appropriées : normalisation par l'osmolalité, normalisation par le *mass spectrometry total useful signal* (MSTUS) et normalisation par la médiane des métabolites par individu rapportés à une référence (les contrôles qualité) (249,250). Nous ne disposons pas de données sur l'osmolalité et la normalisation par MSTUS est utilisée fréquemment dans les études de métabolomique ; nous avons donc procédé à une normalisation par MSTUS pour nos données.

Nous avons exploré les associations entre les expositions disponibles dans la base de données ELISABET et les métabolites urinaires. Pour cela, nous avons fait des régressions linéaires pour tester les associations entre les expositions et les métabolites annotés.

Nous avons dans un premier temps, testé des variables de « démonstration ». Il s'agit de la variable « sexe » car nous nous attendons à un effet du sexe sur l'ensemble du métabolisme, la variable « prise ou non d'aspirine » car nous nous attendons à un impact sur quelques métabolites spécifiques et enfin, la variable « consommation de fruits ». L'étude ELISABET dispose de peu de données sur l'alimentation (seulement une question de fréquence de consommation) mais l'alimentation est connue comme étant fortement associé aux profils

métaboliques (251). Nous nous attendions donc à retrouver des associations même sur une mesure peu précise d'une consommation alimentaire.

Nous avons appliqué une correction des tests multiples par la méthode de Bonferroni et par FDR.

Dans un deuxième temps, nous avons utilisé la même approche pour les variables d'exposition.

Ensuite, nous avons fait une approche par voie métabolique. En effet, notre objectif est de comprendre les mécanismes biologiques sous-jacents, une analyse par voie métabolique en complément d'une analyse par métabolites individuels est recommandée d'après les résultats de la revue systématique.

Pour cela, nous avons utilisé une approche par score qui attribue un score à chaque voie métabolique. Nous avons utilisé le Z-score c'est-à-dire que chaque valeur de métabolite a été transformée en son Z-score, calculé de la manière suivante : (valeur du métabolite – moyenne des métabolites par individu) / écart-type des métabolites par individu (252). Puis nous avons attribué un score à chaque voie métabolique, en faisant la somme des Z-scores des métabolites appartenant à la voie en question.

Un avantage est de pouvoir tester si les métabolites appartenant à une voie, pris dans leur ensemble, sont associés à une exposition (alors que pris individuellement, ils n'ont peut-être pas d'effet).

Nous avons également utilisé la méthode basée sur le rapport de vraisemblance. Nous avons testé le rapport de vraisemblance entre deux modèles statistiques : un premier modèle où la variable d'exposition est placée en variable à expliquer et les variables d'ajustement sont placées parmi les variables explicatives, et un deuxième modèle qui reprend le premier en y ajoutant le score de la voie métabolique dans les variables explicatives.

Le rapport de vraisemblance teste la contribution des voies métaboliques dans le modèle c'est à dire qu'on teste si le modèle incluant les scores de voies métaboliques permet de mieux « prédire » la variable d'exposition que les autres variables d'ajustement seules.

Les représentations graphiques utilisées sont le volcano plot et le diagramme quantile-quantile (QQ plot).

Sur le volcano plot sont représentés, en abscisse le coefficient β de l'association et en ordonnées $-\log_{10}$ pvalue. Chaque point représente un métabolite. Le seuil de significativité est calculé en fonction du nombre de tests réalisés (correction de Bonferroni ou de Benjamini-Hochberg) et est représenté par une ligne horizontale pointillée.

Si le QQ plot montre une déviation de l'attendu alors on pourra regarder le test du rapport de vraisemblance réalisé sur l'ensemble des voies métaboliques. Pour cela, nous utilisons deux modèles : un premier modèle où la variable d'exposition est placée en variable à expliquer et les variables d'ajustement sont placées parmi les variables explicatives, et un deuxième modèle qui reprend le premier en y ajoutant tous les scores des voies métaboliques dans les variables explicatives. Si le test du rapport de vraisemblance est significatif alors la déviation pourrait être liée à une association avec l'exposition sur les métabolites ou les voies métaboliques ; sinon il faut plutôt conclure à un effet lié à une structure de population ou corrélation entre les différentes variables attendues (à noter que la corrélation entre métabolites peut avoir un sens physiologique, et la variation sur un seul processus

métabolique peut impacter mécaniquement de nombreux métabolites impliqués en aval de cette réaction). Ce test n'est réalisable que si le nombre de variables est limité et très inférieur au nombre de sujets dans l'étude. Ce test est donc pertinent pour les voies métaboliques qui comportent, en métabolomique, généralement un nombre limité de métabolites annotés dans la base de données. Un test de rapport de vraisemblance significatif indique qu'au moins une des variables étudiées (voies métaboliques) est associée à l'exposition. Ce test a l'avantage de répondre à cette question avec un seul test statistique limitant le nombre de tests. L'inconvénient est qu'il porte sur un nombre de degrés de liberté égal au nombre de variables traitées, ce qui réduit sa puissance. De plus, il ne permet pas directement d'identifier les variables associées à l'exposition. Des analyses complémentaires doivent être réalisées.

Les analyses sur l'ensemble des métabolites sont présentées en **annexe 3**.

3. Résultats

a) Population

Nous avons sélectionné aléatoirement 1 000 individus issus de l'étude ELISABET. Parmi eux, 27 ont été exclus en raison d'un retrait de consentement, à la suite de l'envoi d'une lettre d'information dans le cadre de la conformité au RGPD (règlement général sur la protection des données) demandée par le promoteur de l'étude. De plus, 47 individus présentaient un échantillon d'urine non exploitable pour les analyses de chromatographie spectrométrie de masse. Au total, 926 individus ont finalement été inclus dans cette étude.

Le tableau III.D.3.1 présente une description des participants.

Tableau III.D.3.1. Caractéristiques des participants de l'étude

	N = 926
Age (années) (moyenne (ET))	53,32 (7,28)
Sexe (hommes) (n (%))	404 (43,63)
Communauté urbaine (n (%))	
Lille	472 (50,97)
Dunkerque	454 (49,03)
Prise d'aspirine (n (%))	36 (3,89)
Consommation de fruits (nombre de prises par jour) (n (%))	
0	17 (1,84)
0,033	14 (1,51)
0,066	15 (1,62)
0,14	44 (4,75)
0,36	126 (13,61)
0,714	74 (7,99)
1	332 (35,85)
2	210 (22,68)
3	94 (10,15)
Polluants	
PM ₁₀ court terme (µg/m ³) (médiane (IIQ))	22,04 [16,83-34,17]
PM ₁₀ long terme (µg/m ³) (médiane (IIQ))	26,63 [25,66-27,54]
NO ₂ court terme (µg/m ³) (médiane (IIQ))	21,88 [16,25-31,44]
NO ₂ long terme (µg/m ³) (médiane (IIQ))	21,64 [19,06-25,75]
O ₃ court terme (µg/m ³), médiane [IIQ]	60,20 [44,05-71,90]
Moisissures, oui (n (%))	192 (20,73)
Vapeur (score arbitraire) (médiane (IIQ))	0 [0,00-2,75]
Poussières (score arbitraire) (médiane (IIQ))	0 [0,00-5,75]
Gaz (score arbitraire) (médiane (IIQ))	0 [0,00-0,00]
Fumées (score arbitraire) (médiane (IIQ))	0 [0,00-0,00]
Cadmium (µg/L) (médiane (IIQ))	0,27 [0,17- 0,39]
Plomb (µg/L) (médiane (IIQ))	17,70 [12,20 - 25,40]

Abréviations : ET = écart type ; PM₁₀ = particules en suspension ayant un diamètre aérodynamique inférieur à 10 µm ; IIQ = intervalle interquartile ; NO₂ = dioxyde d'azote ; O₃ = ozone

b) Variables de démonstrations et expositions

Concernant les variables de démonstration, le sexe est une variable binaire (homme ou femme), la prise d'aspirine est une variable binaire (oui ou non), la consommation de fruits est une variable quantitative correspondant à la fréquence de consommation de fruits d'après les réponses des participants.

Les méthodes de mesure des expositions environnementales ont été décrites dans la partie Méthode générale et les expositions sont décrites dans le tableau III.D.3.1.

c) Métabolites

Les analyses chimiques ont permis la mesure de 10 043 signaux. Parmi eux, certains ont conduit à des annotations de métabolites : 327 présentaient un niveau de preuve « a,b,c » (identification confirmée) ou « a,b » (identification putative). Au total, 4 331 avaient un niveau de preuve « a » et 5 385 n'avaient aucun niveau de preuve. Pour ces deux dernières catégories, l'identification des métabolites n'était pas possible à ce stade.

Les métabolites annotés appartenaient à 74 voies métaboliques différentes, chacune comprenant entre 1 et 21 métabolites (Voir le tableau en **annexe 4**).

d) Analyses

(1) Preuve de concept : variables de démonstration

Les analyses sur la consommation de fruits ont montré de nombreux métabolites et de nombreuses voies métaboliques associés à la consommation de fruits (figures III.D.3.1, III.D.3.2, III.D.3.3, III.D.3.4 et III.D.3.5). Il était attendu que la consommation de fruits, et, plus largement, le régime alimentaire soient associés au profil métabolique (253).

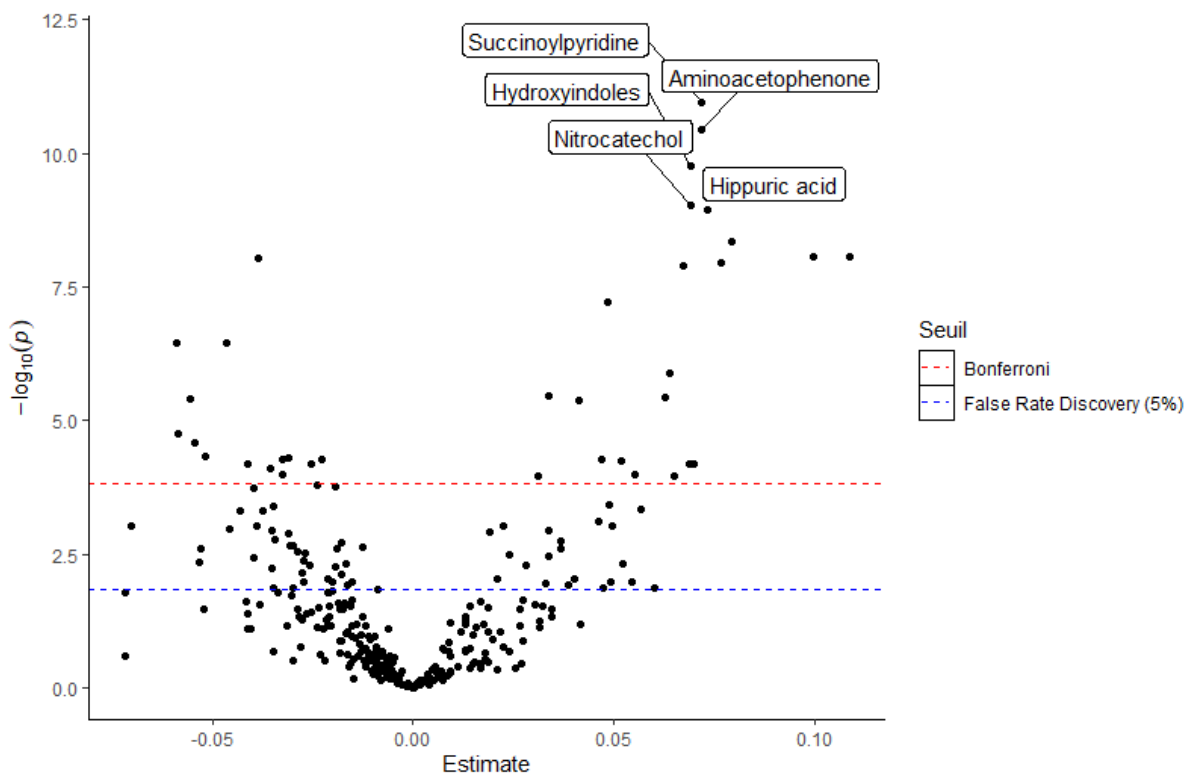


Figure III.D.3.1 : Volcano plot représentant les associations entre la consommation de fruits et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

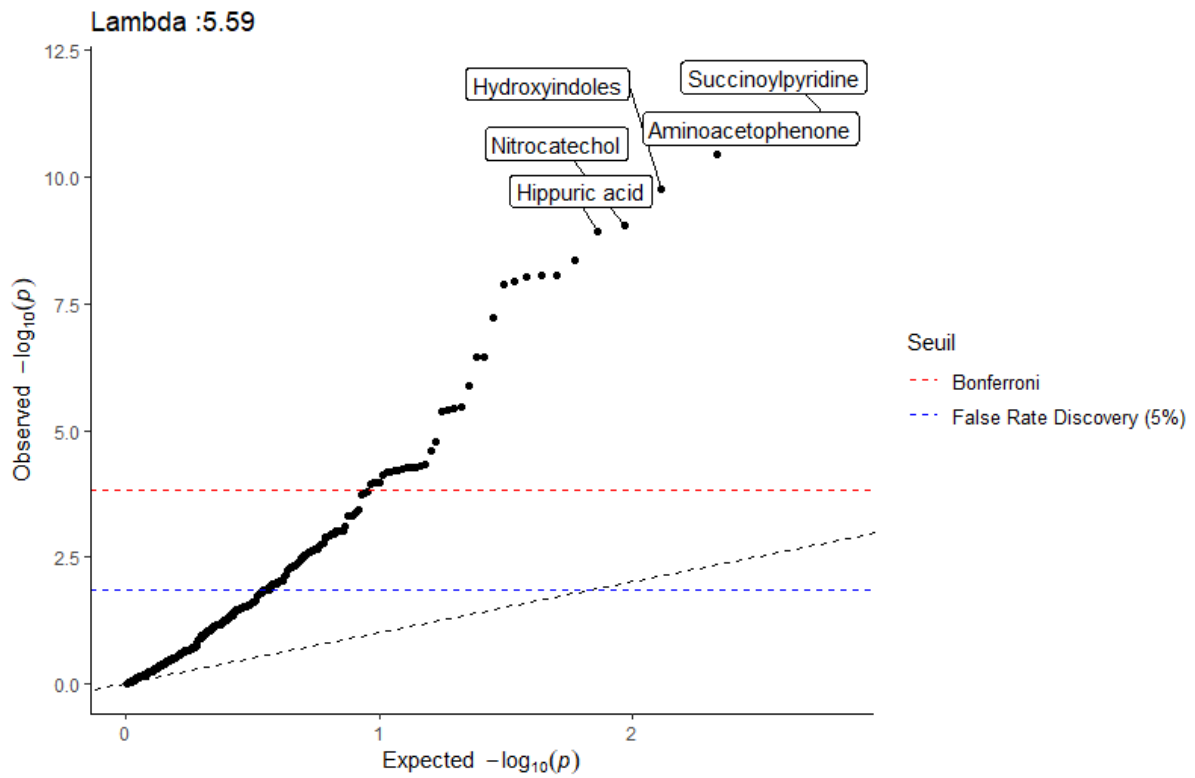


Figure III.D.3.2 : QQ plot représentant les associations entre la consommation de fruits et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

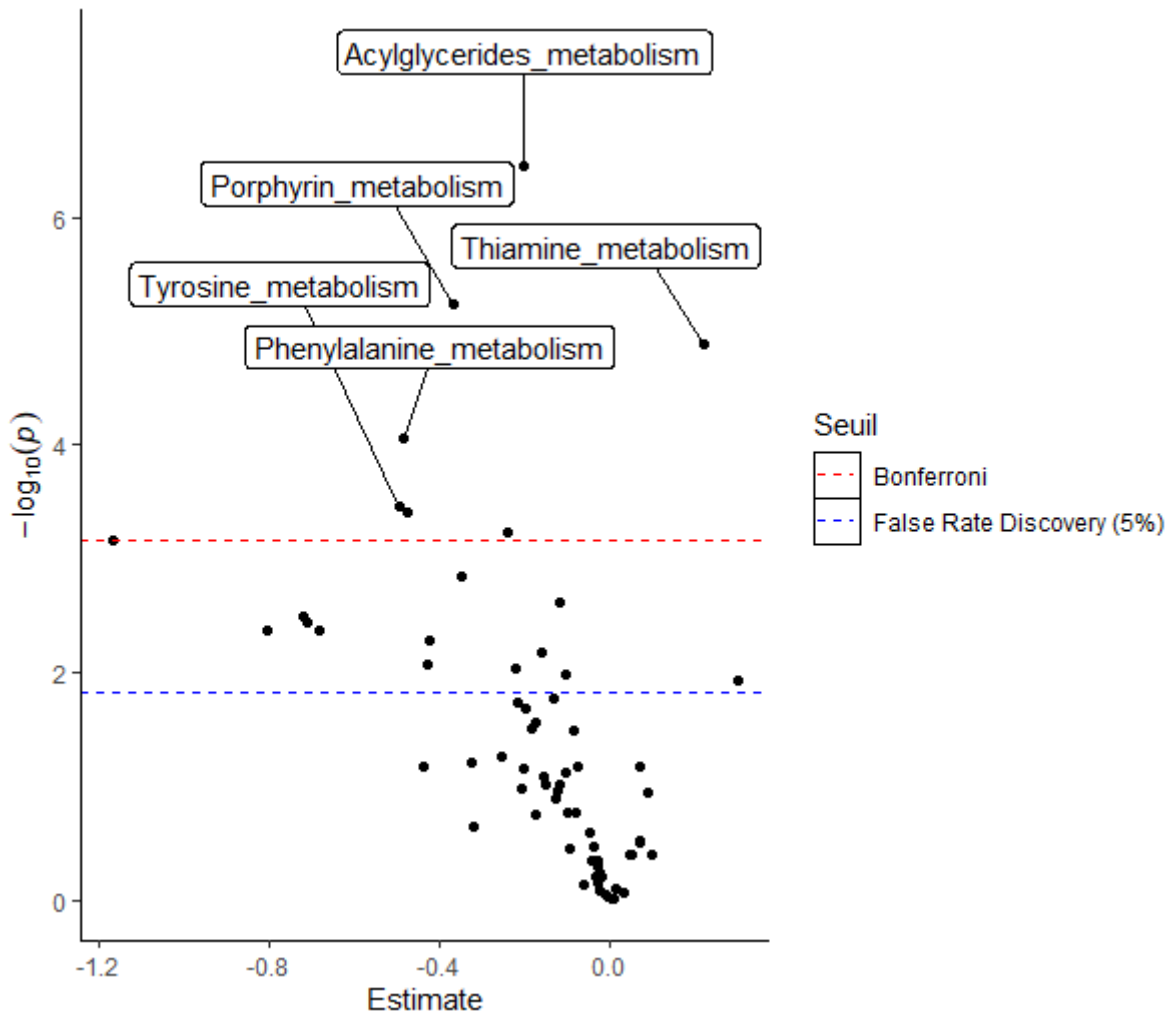


Figure III.D.3.3 : Volcano plot représentant les associations entre la consommation de fruits et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

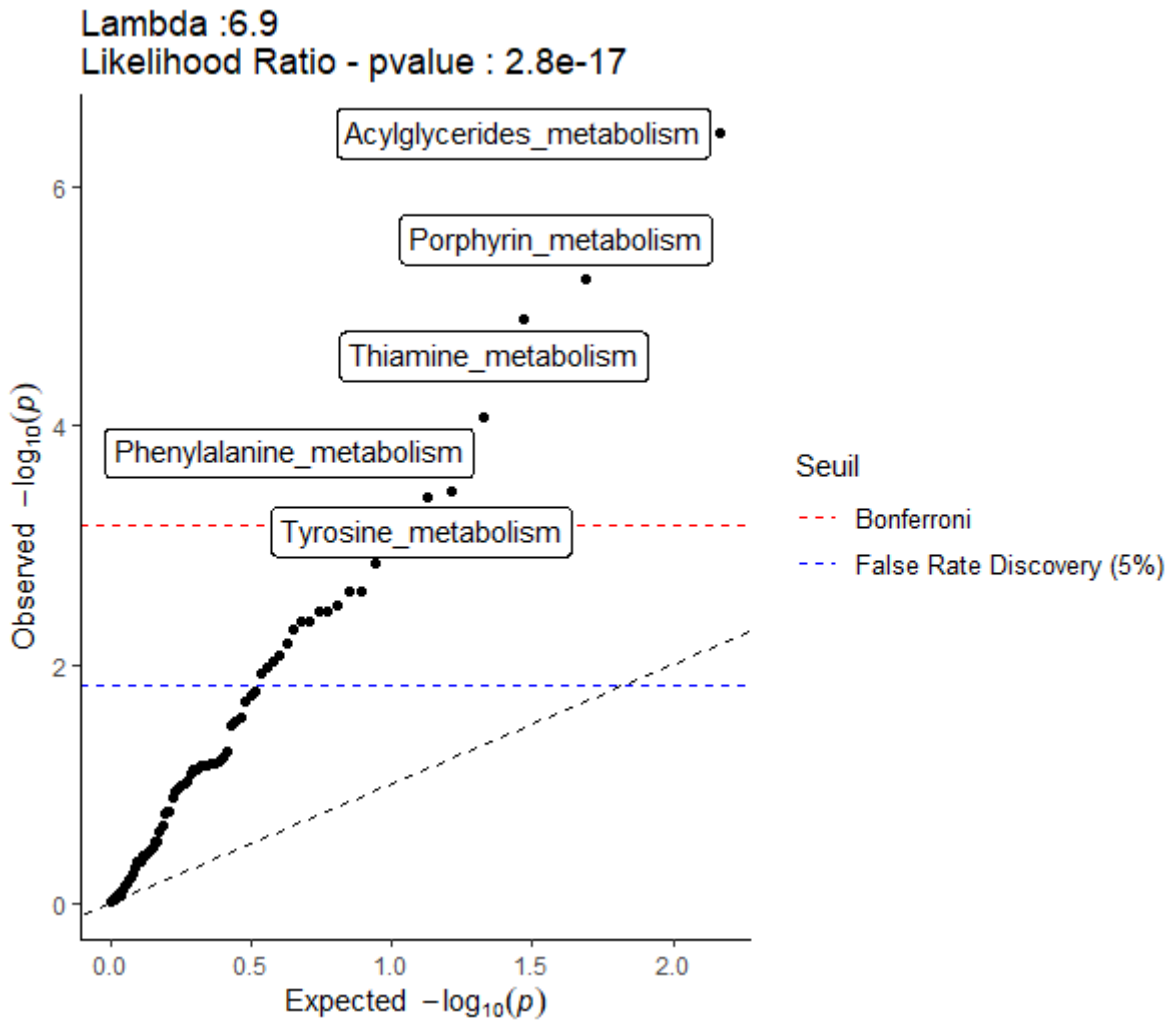


Figure III.D.3.4 : QQ plot représentant les associations entre la consommation de fruits et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

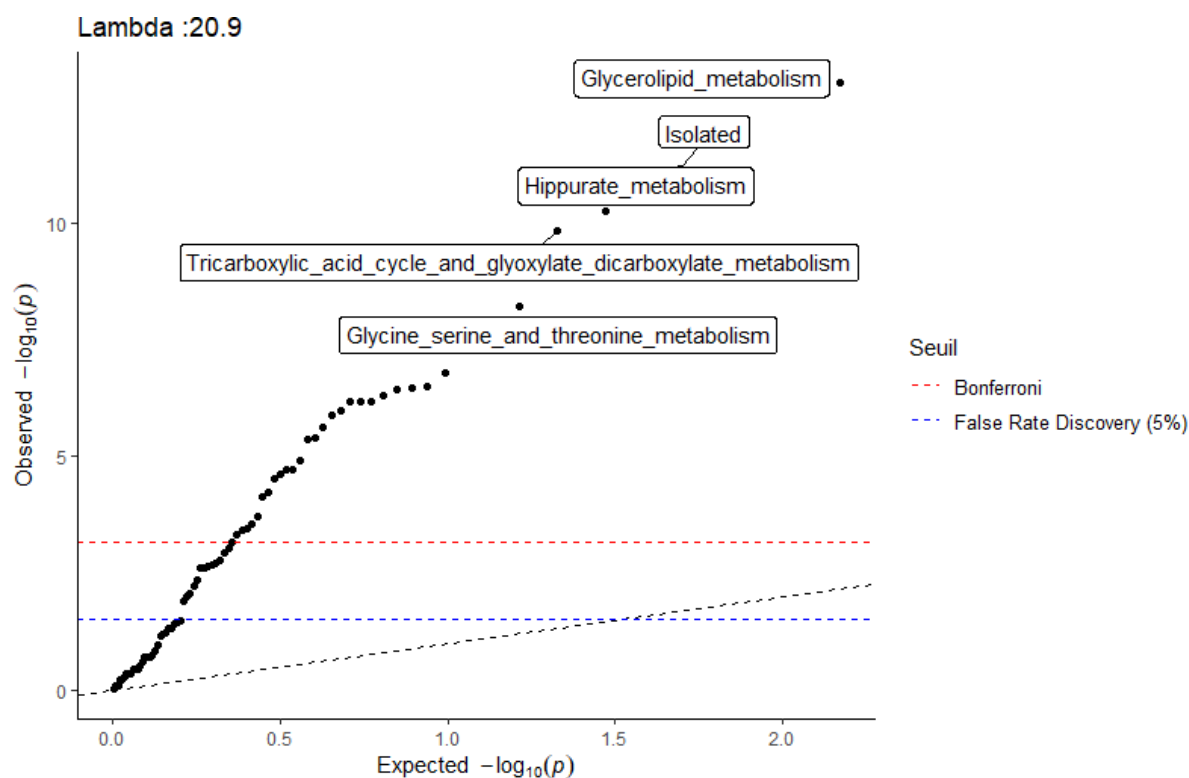


Figure III.D.3.5 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre la consommation de fruits et les voies métaboliques (Z-score) ajustées sur l'âge, le sexe et la communauté urbaine

Les autres variables de démonstration (sexe et prise d'aspirine) ont été testées. Nous avons retrouvé de multiples associations entre le sexe (hommes comparés aux femmes) et les métabolites ainsi que les voies métaboliques (Z-score). Ces résultats suggèrent beaucoup de différences entre le métabolisme des hommes et celui des femmes. Ces différences entre les hommes et les femmes étaient attendues (254).

Ensuite, nous avons regardé la variable « aspirine » (prise d'aspirine comparée à l'absence de prise d'aspirine) et nous avons retrouvé 3 métabolites significativement augmentés avec la prise d'aspirine : l'acide o-hydroxyhippurique, la lysine et le dihydroxybenzaldéhyde. L'acide o-hydroxyhippurique est un métabolite connu de l'aspirine (255), la lysine est un acide aminé retrouvé dans un certains médicaments contenant de l'acide acétylsalicylique, comme le Kardégic (256). Nous n'avons pas trouvé de lien évident dans la littérature entre l'aspirine et le dihydroxybenzaldéhyde. Ces résultats sont cohérents, puisque l'on ne s'attend pas à avoir un impact sur l'ensemble des métabolites avec la prise d'un médicament mais bien sur quelques métabolites spécifiques. Cependant, lorsqu'il y a de nombreuses voies métaboliques impactées, il est complexe de toutes les analyser et surtout d'identifier lesquelles sont les plus impactées. Ce point est développé dans la Discussion générale.

(2) Pollution atmosphérique

Nous avons étudié les associations entre l'exposition aux PM₁₀, NO₂ et O₃ et les métabolites ainsi que les voies métaboliques.

Les résultats pour l'exposition aux PM₁₀ à court terme ont montré une association avec un métabolite après correction des tests multiples par FDR : l'acide 3,4 dihydroxybenzoïque (figures III.D.3.6 et III.D.3.7). Celui-ci est un composé ayant des propriétés antioxydantes (257), et est retrouvé dans certains aliments (par exemple dans le thé) (258). Une hypothèse serait que la diminution observée de ce composé s'explique par une altération du métabolisme oxydatif provoquée par les PM₁₀ (259). En revanche, aucune voie métabolique n'était associée aux PM₁₀ et le rapport de vraisemblance sur l'ensemble des voies métaboliques (Z-score) n'était pas significatif (figures III.D.3.8, III.D.3.9 et III.D.3.10).

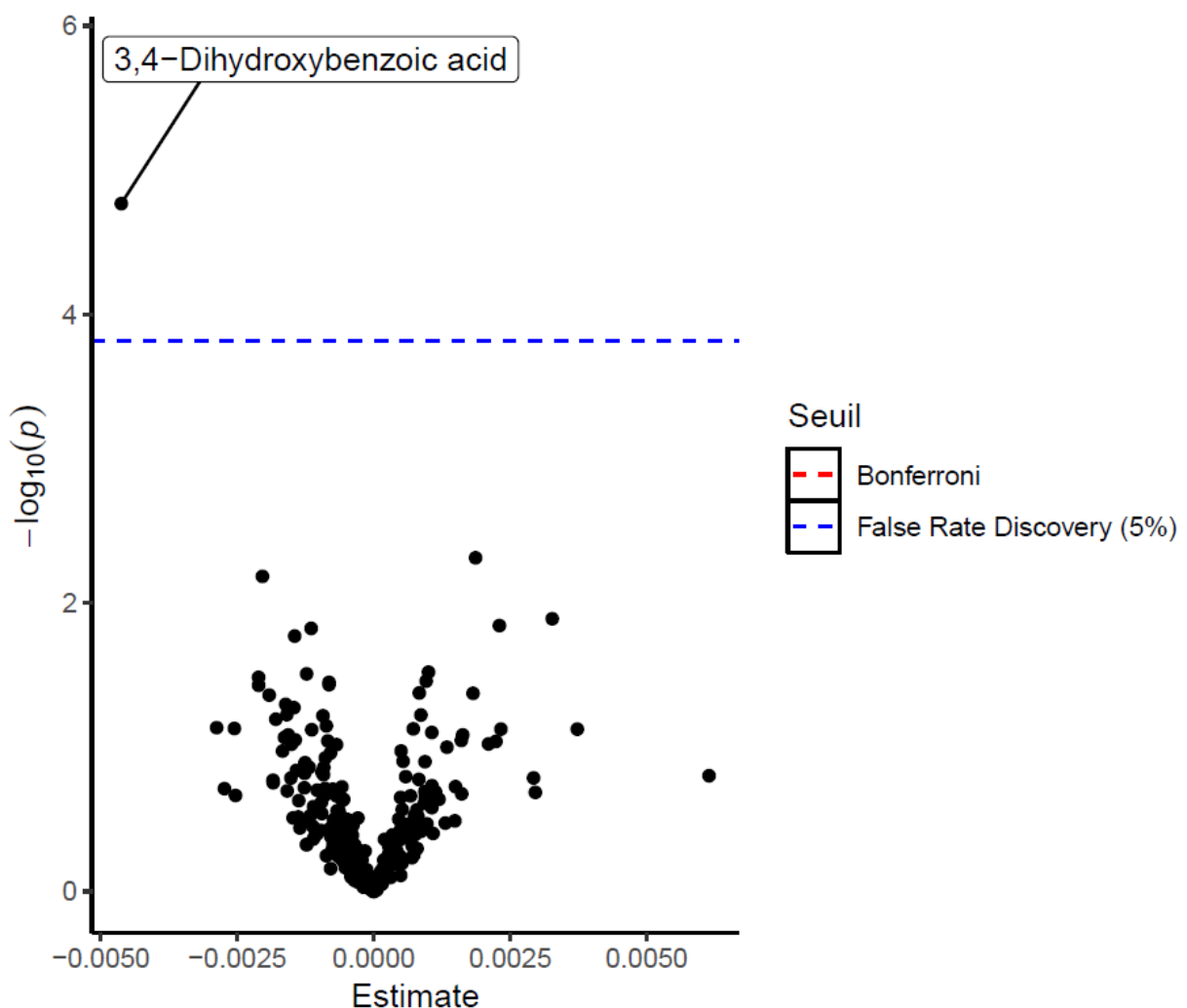


Figure III.D.3.6 : Volcano plot représentant les associations entre l'exposition à court terme aux PM₁₀ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

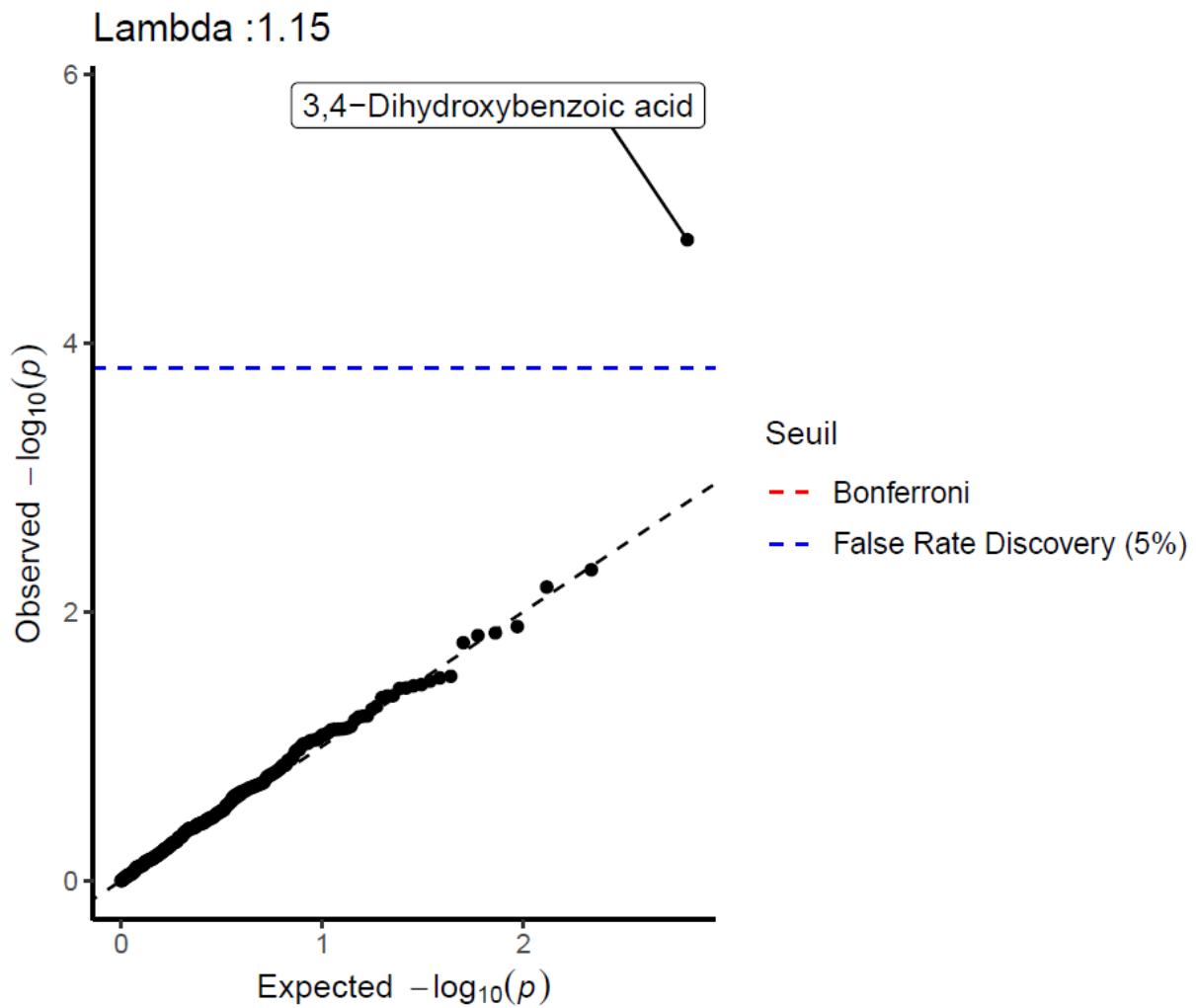


Figure III.D.3.7 : QQ plot représentant les associations entre l'exposition à court terme aux PM_{10} et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

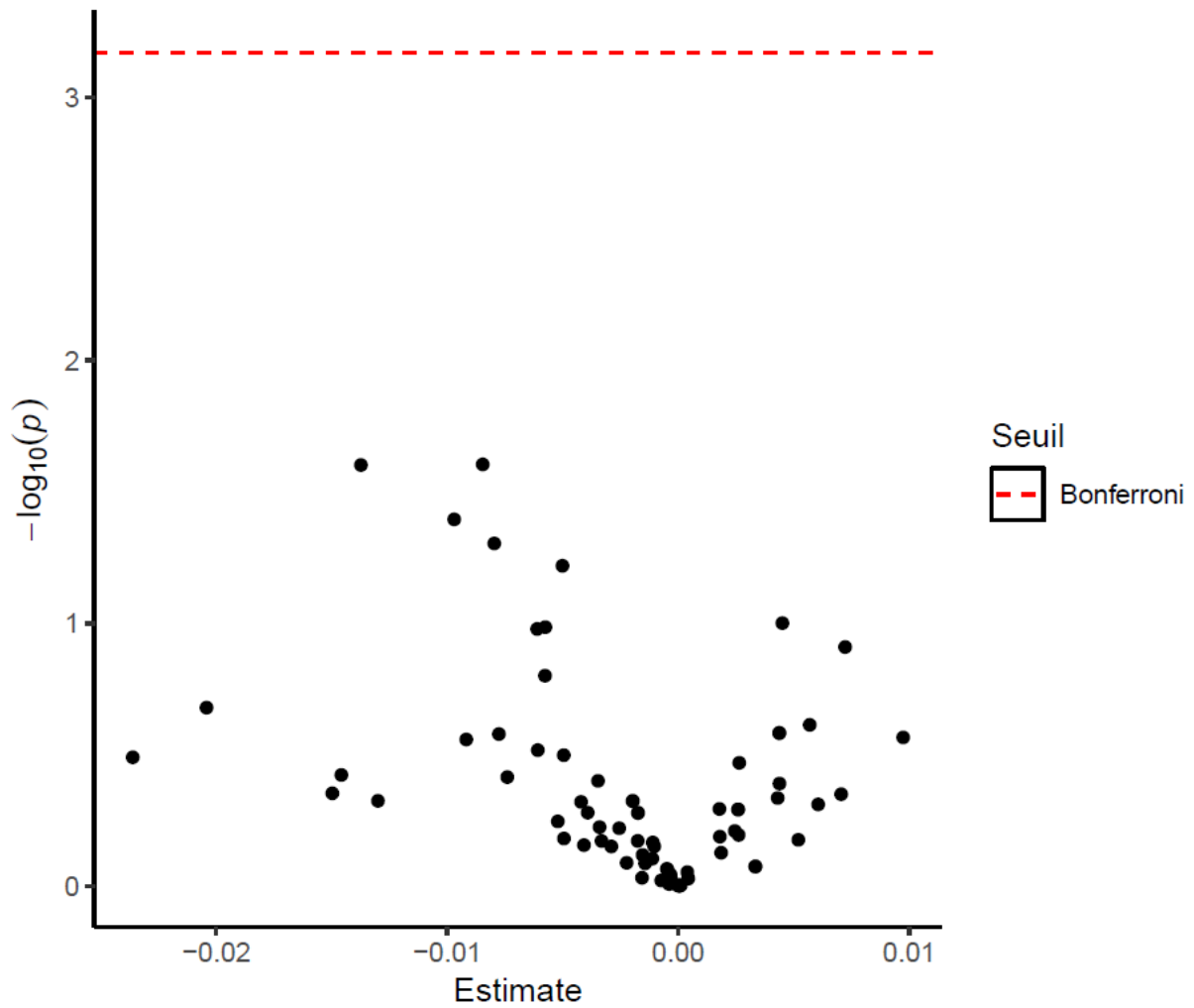


Figure III.D.3.8 : Volcano plot représentant les associations entre l'exposition à court terme aux PM_{10} et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

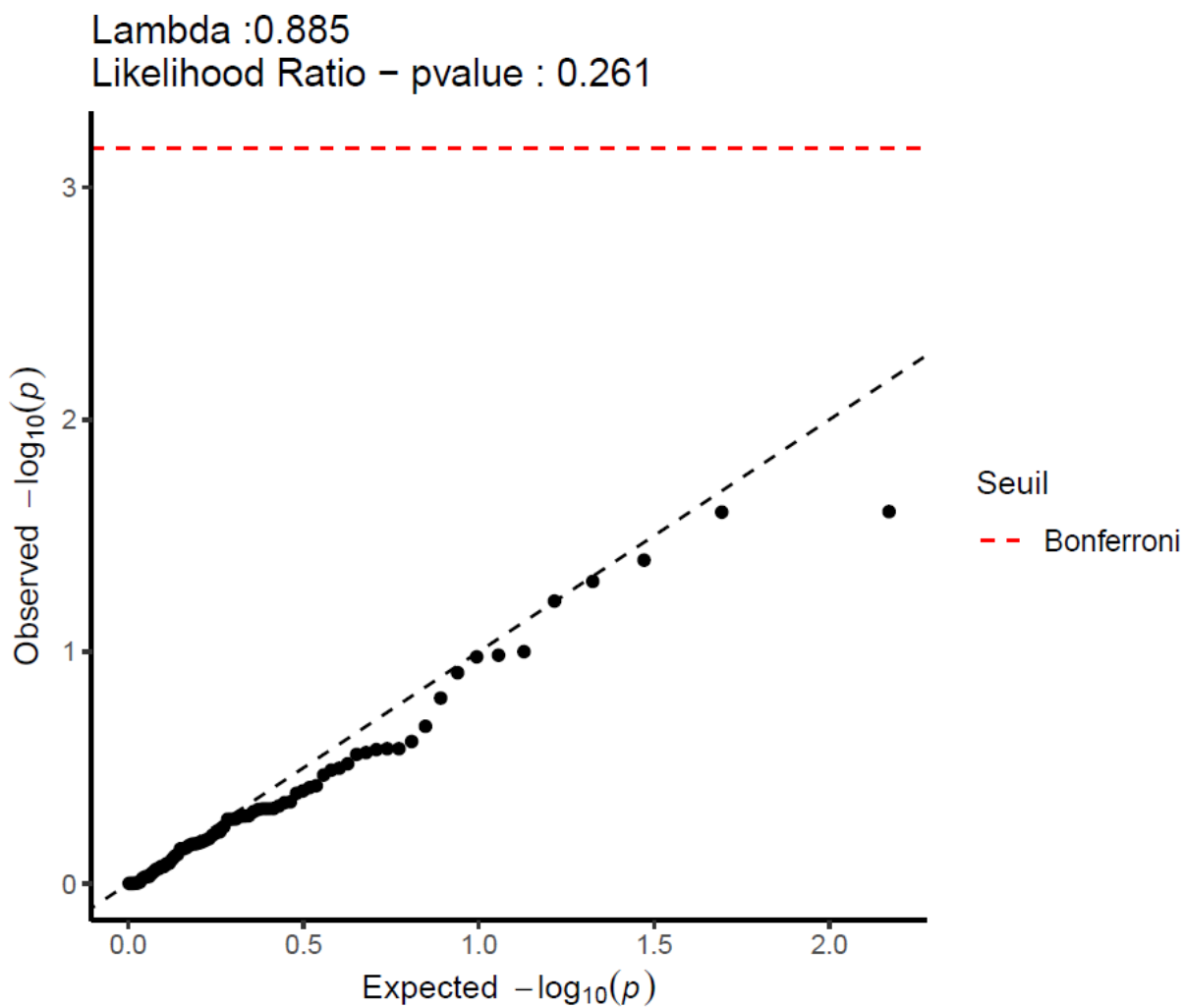


Figure III.D.3.9 : QQ plot représentant les associations entre l'exposition à court terme aux PM_{10} et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

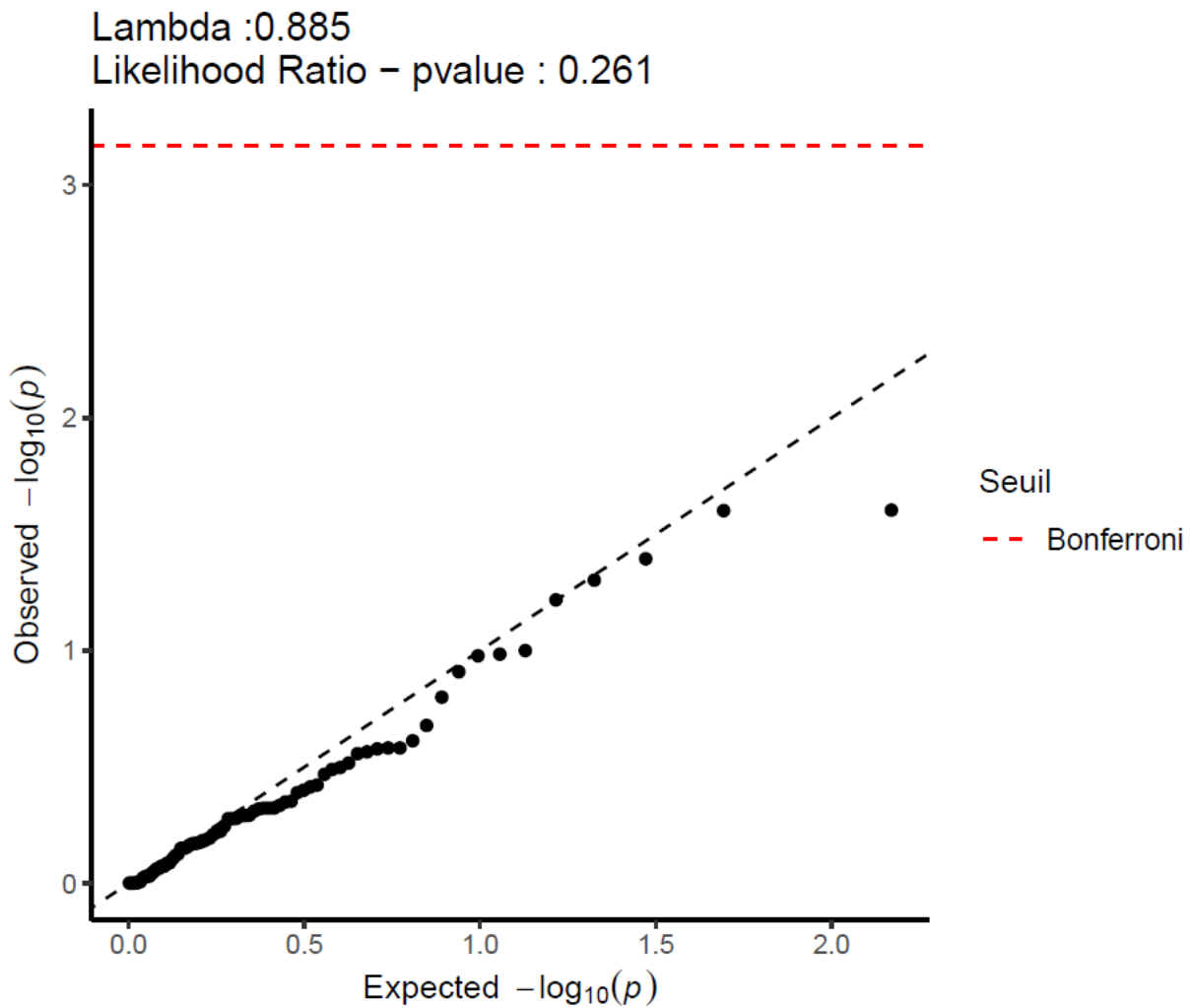


Figure III.D.3.10 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre l'exposition à court terme aux PM_{10} et les voies métaboliques (Z score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

Concernant l'exposition résidentielle aux PM_{10} , aucune association n'était retrouvée ni avec les métabolites, ni avec les voies métaboliques (Z-score) (figures III.D.3.11, III.D.3.12, III.D.3.13, III.D.3.14 et III.D.3.15).

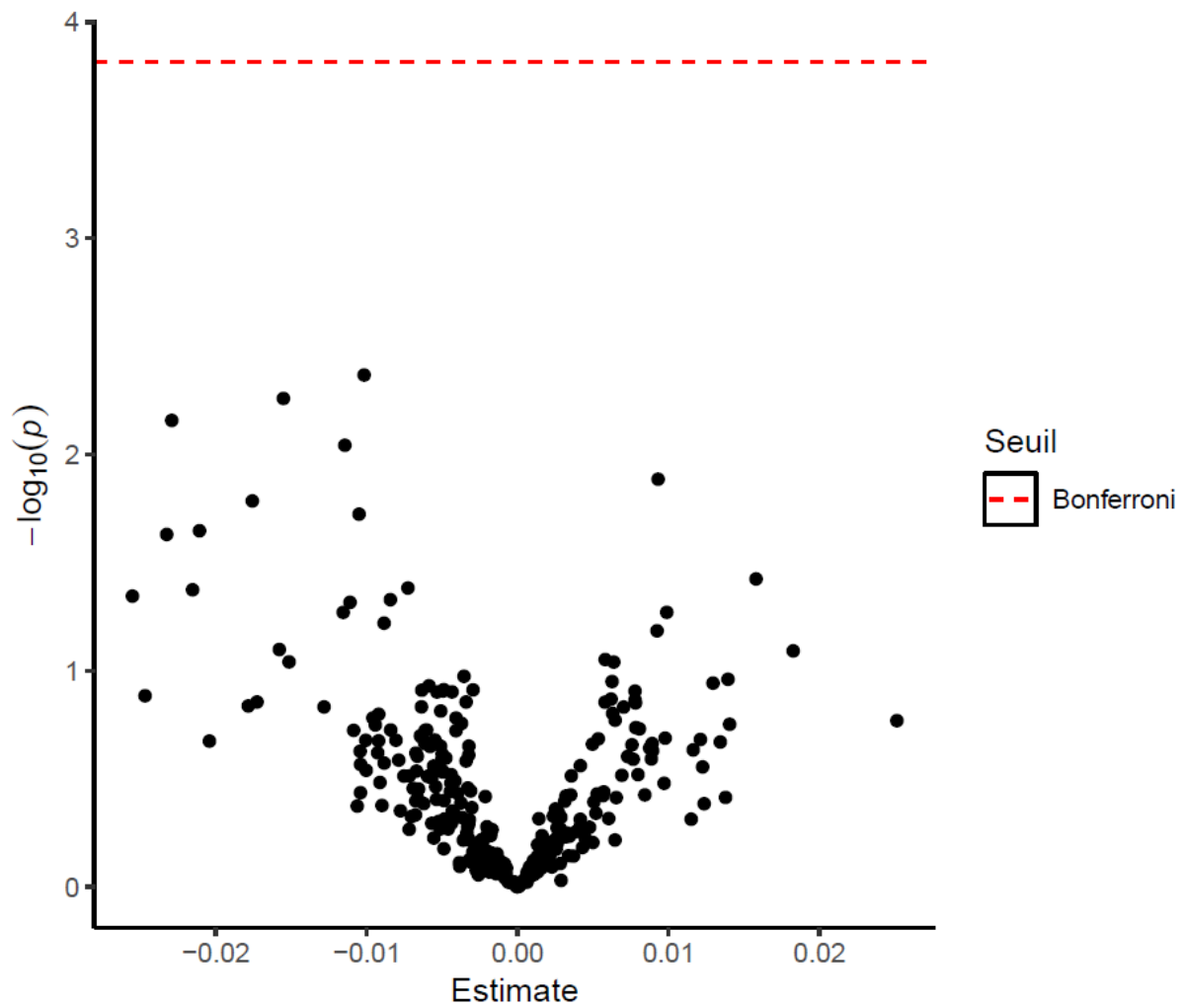


Figure III.D.3.11 : Volcano plot représentant les associations entre l'exposition résidentielle aux PM₁₀ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

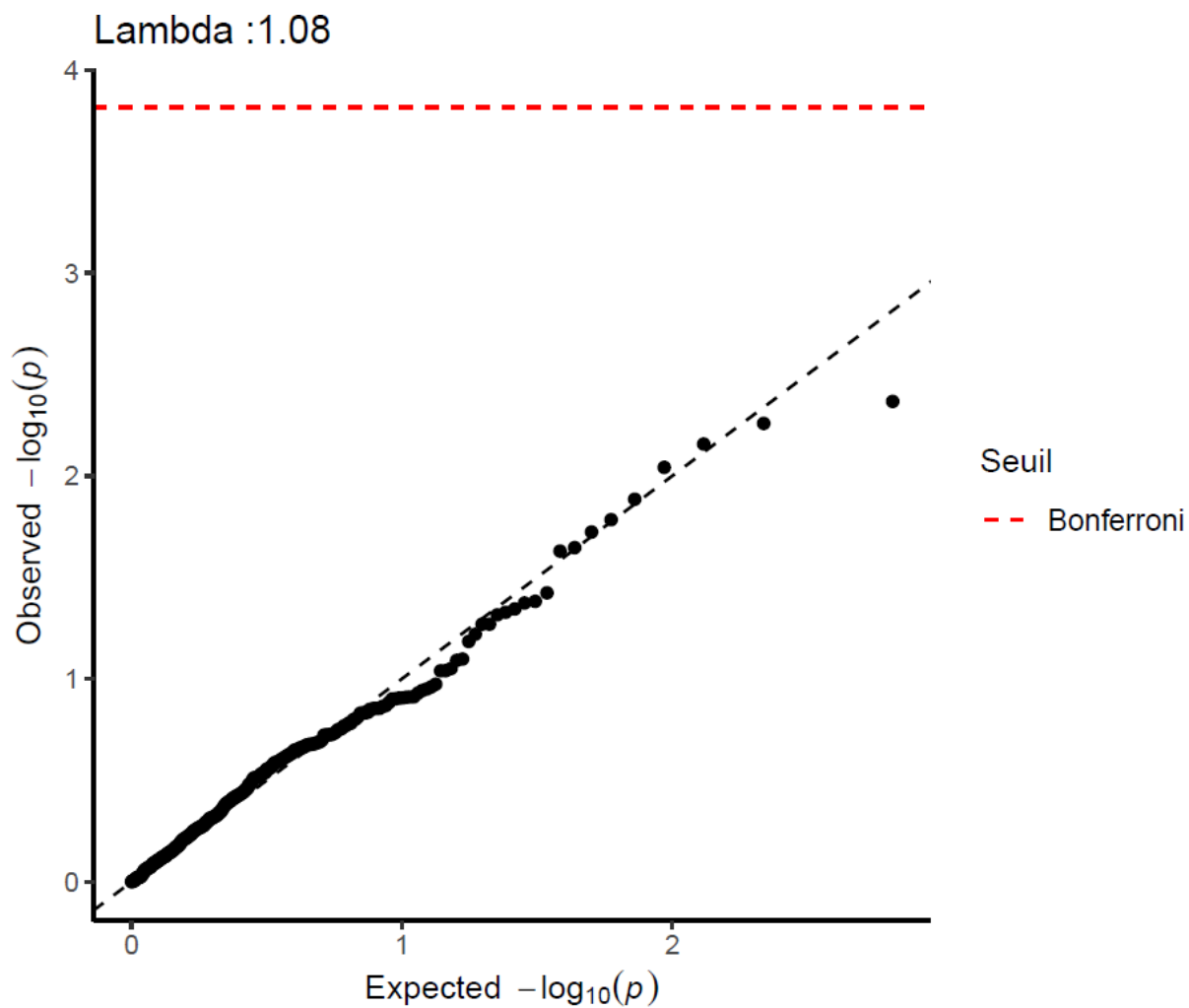


Figure III.D.3.12 : QQ plot représentant les associations entre l'exposition résidentielle aux PM_{10} et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

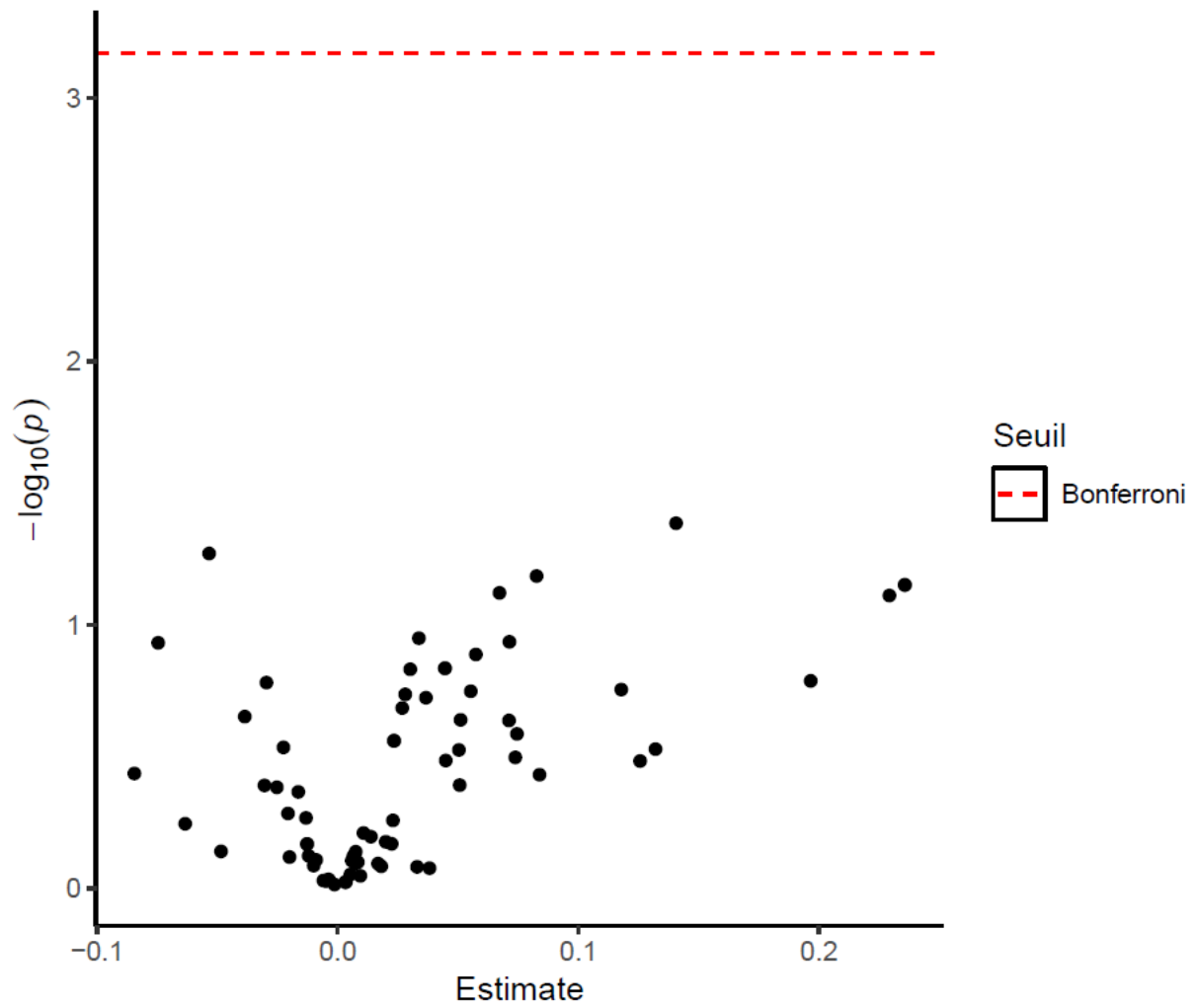


Figure III.D.3.13 : Volcano plot représentant les associations entre l'exposition résidentielle aux PM₁₀ et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

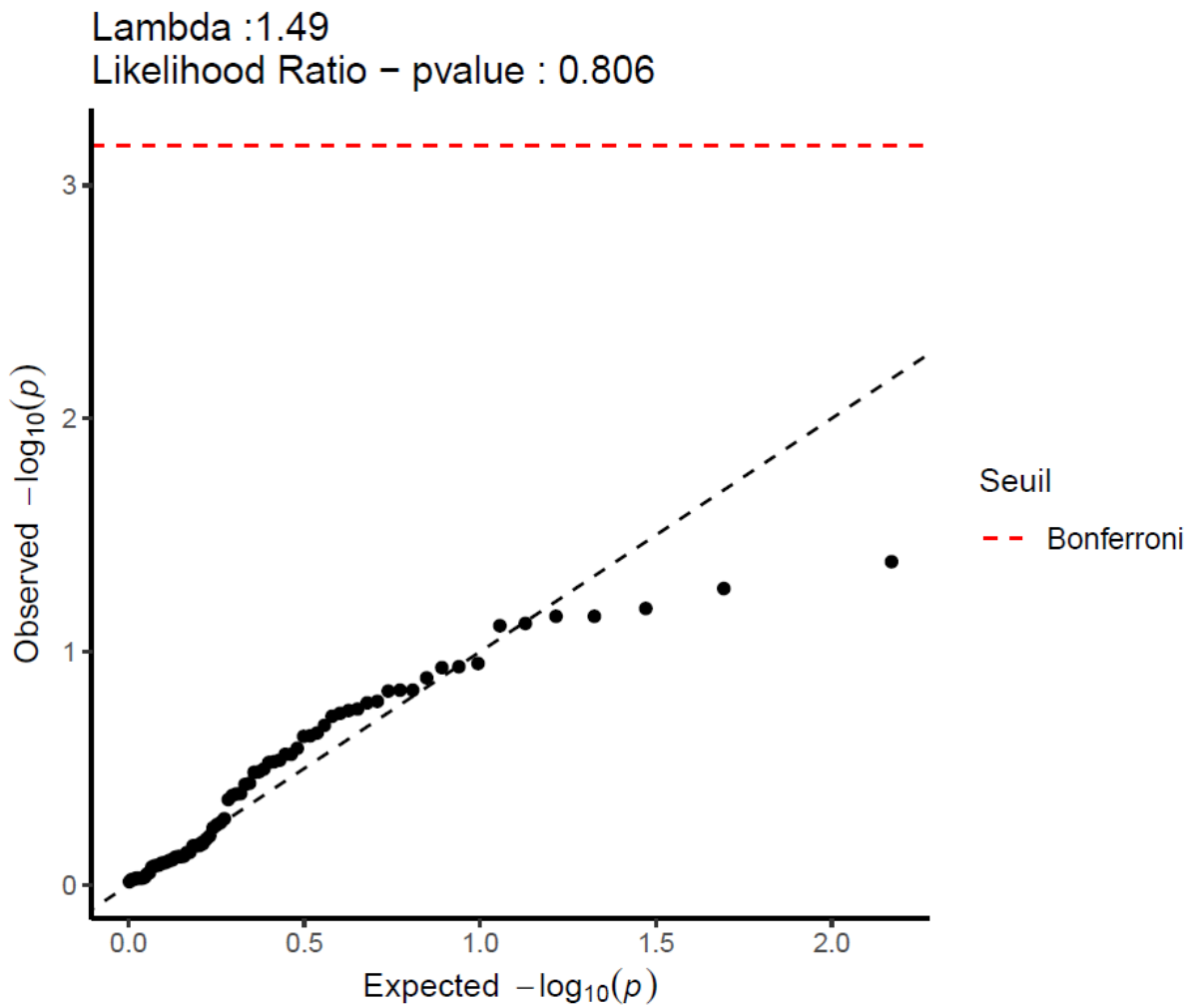


Figure III.D.3.14 : QQ plot représentant les associations entre l'exposition résidentielle aux PM_{10} et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

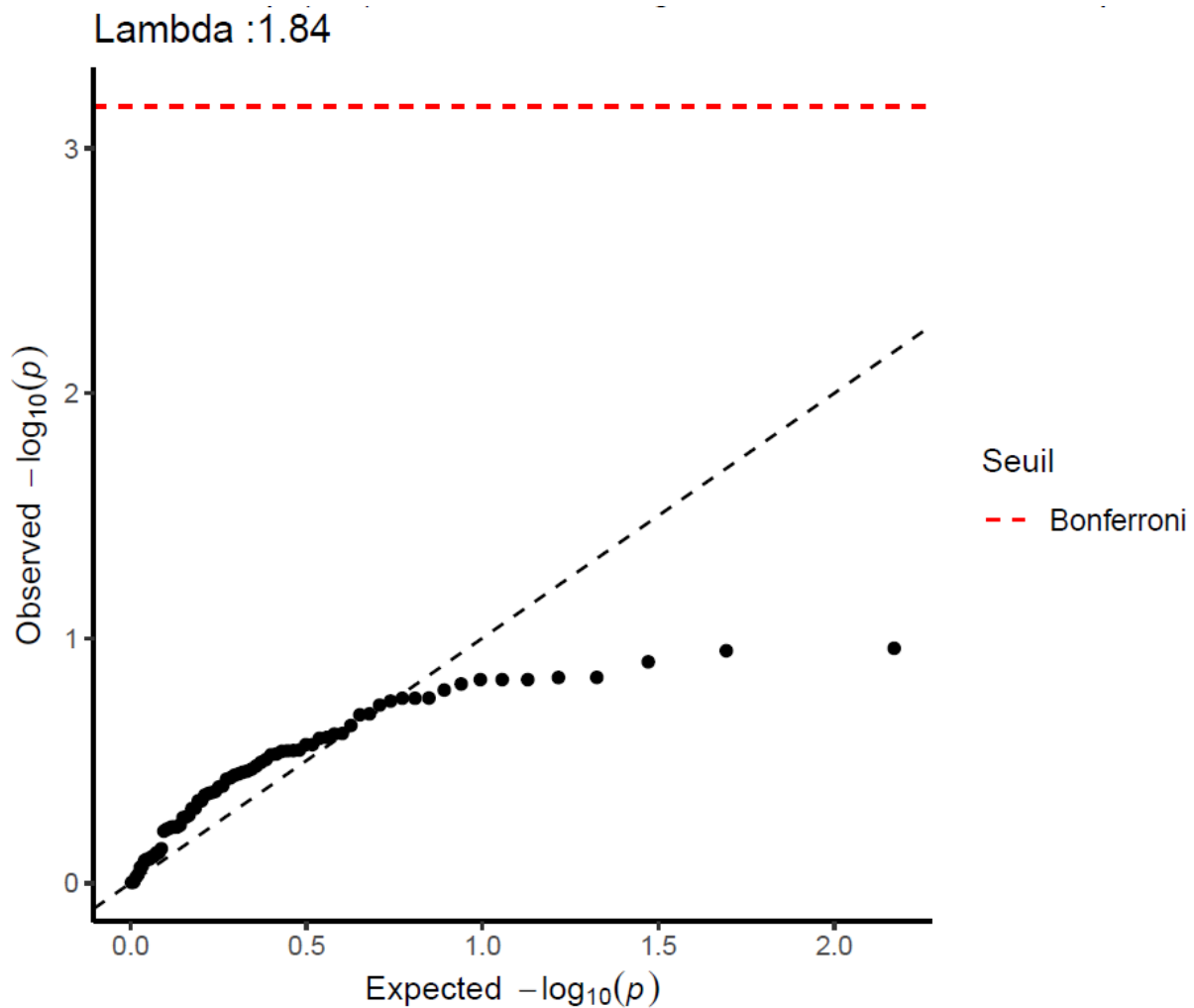


Figure III.D.3.15 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre l'exposition résidentielle aux PM_{10} et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

Concernant les expositions à court terme (figures III.D.3.16, III.D.3.17, III.D.3.18, III.D.3.19 et III.D.3.20) et résidentielle au NO_2 (figures III.D.3.21, III.D.3.22, III.D.3.23, III.D.3.24 et III.D.3.25), ainsi que l'exposition à court terme à l' O_3 (figures III.D.3.26, III.D.3.27, III.D.3.28, III.D.3.29 et III.D.3.30), aucune association n'était significative.

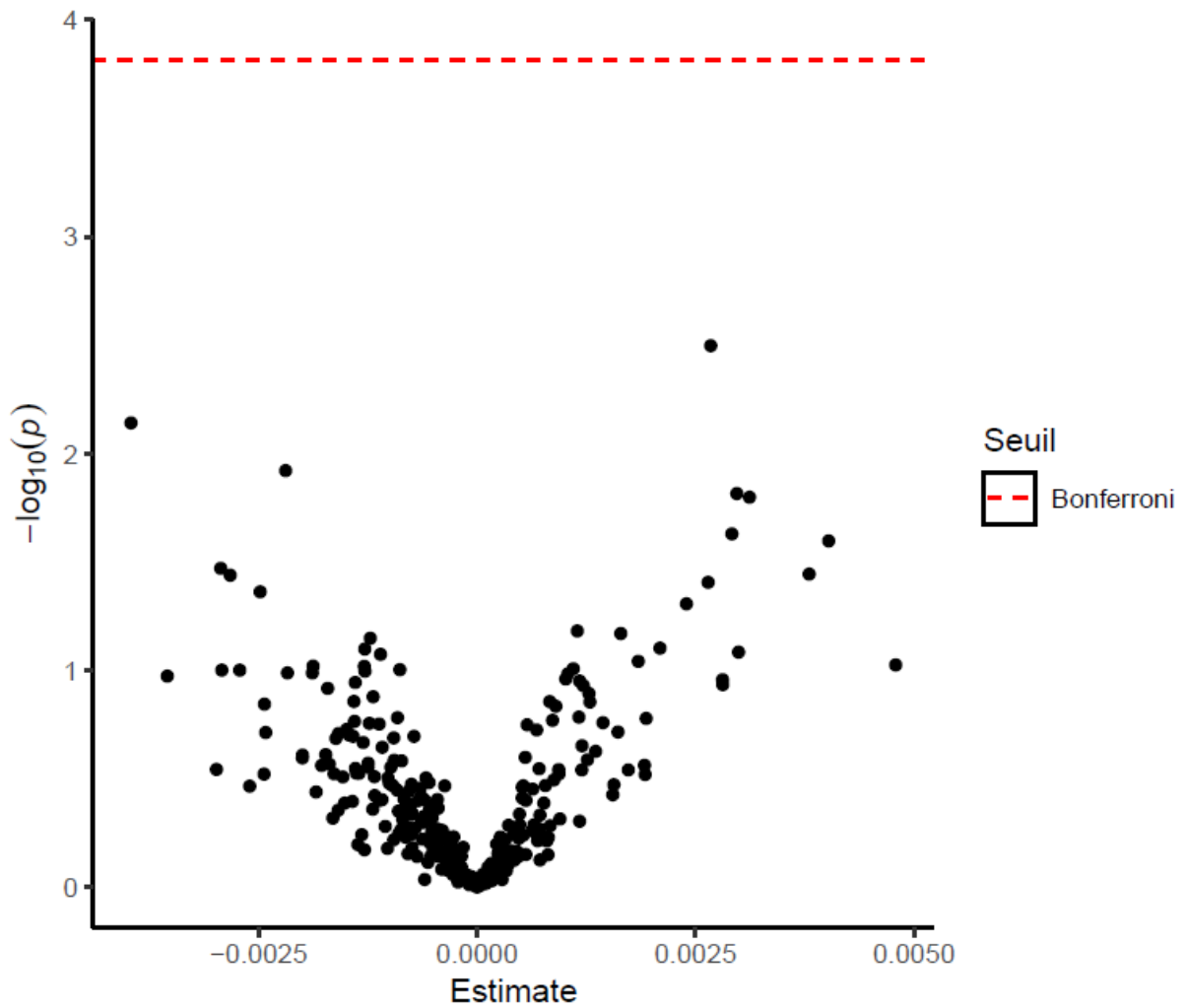


Figure III.D.3.16 : Volcano plot représentant les associations entre l'exposition à court terme au NO₂ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

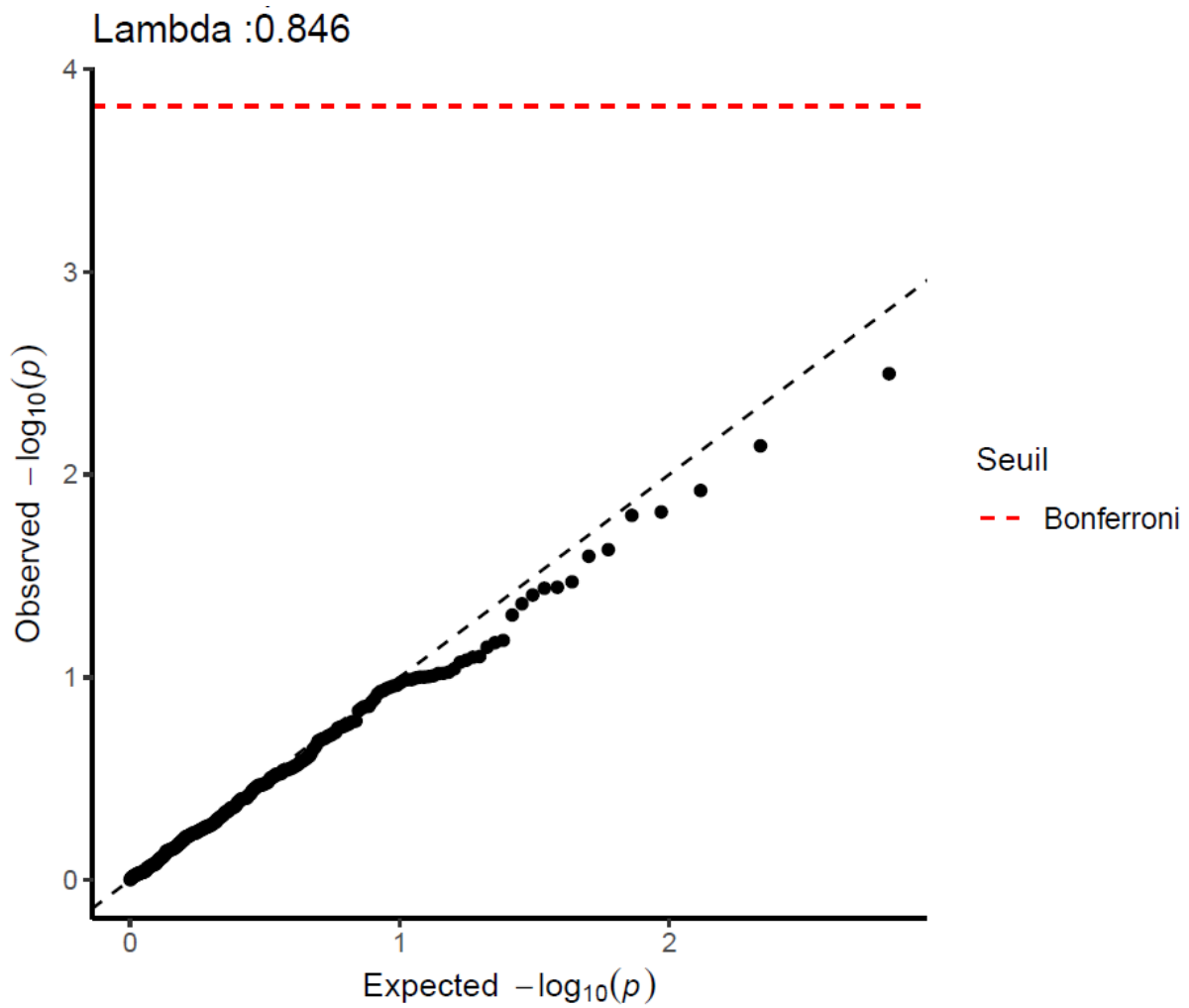


Figure III.D.3.17 : QQ plot représentant les associations entre l'exposition à court terme au NO_2 et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

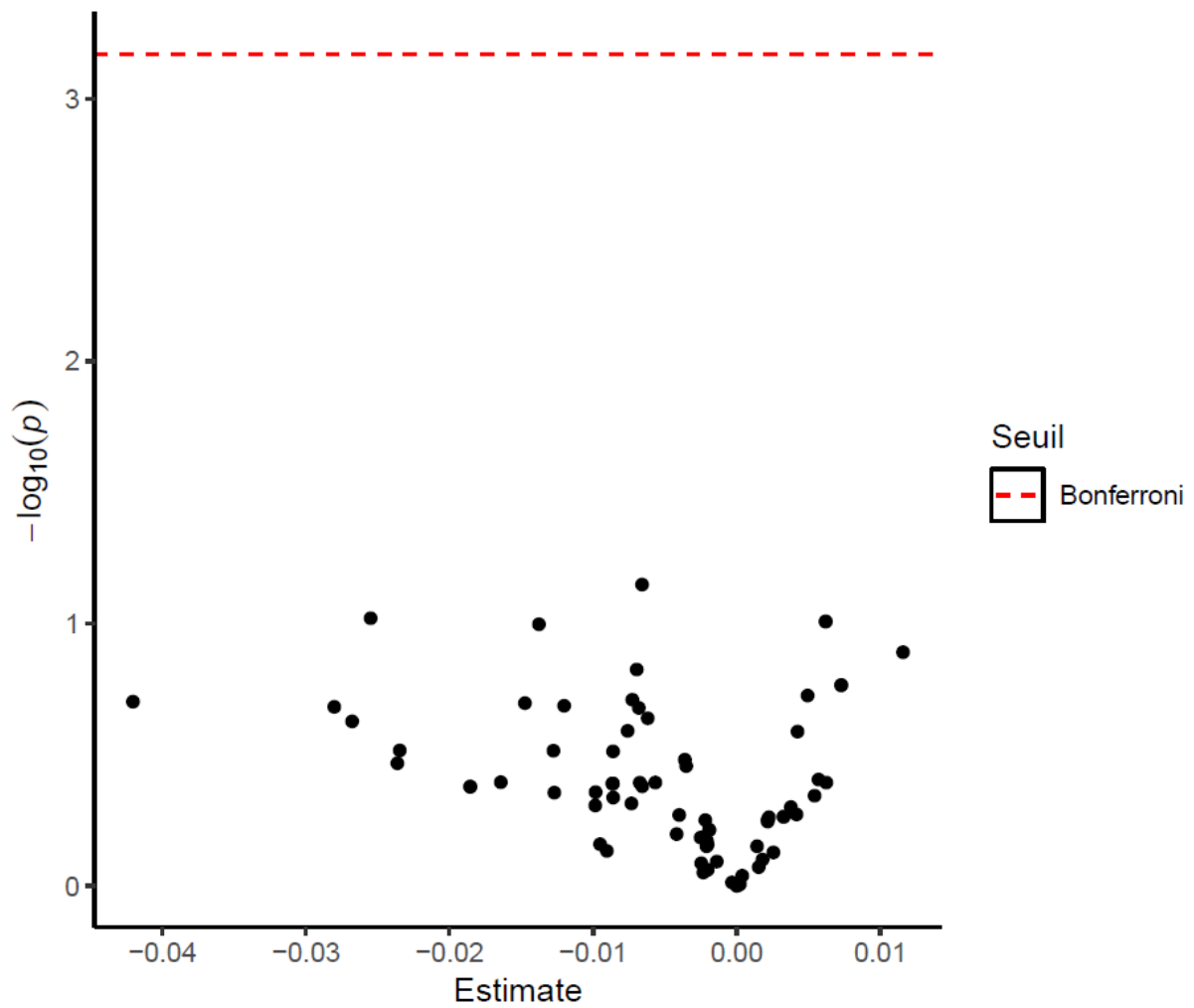


Figure III.D.3.18 : Volcano plot représentant les associations entre l'exposition à court terme au NO₂ et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

Lambda :1.38
Likelihood Ratio – pvalue : 0.825

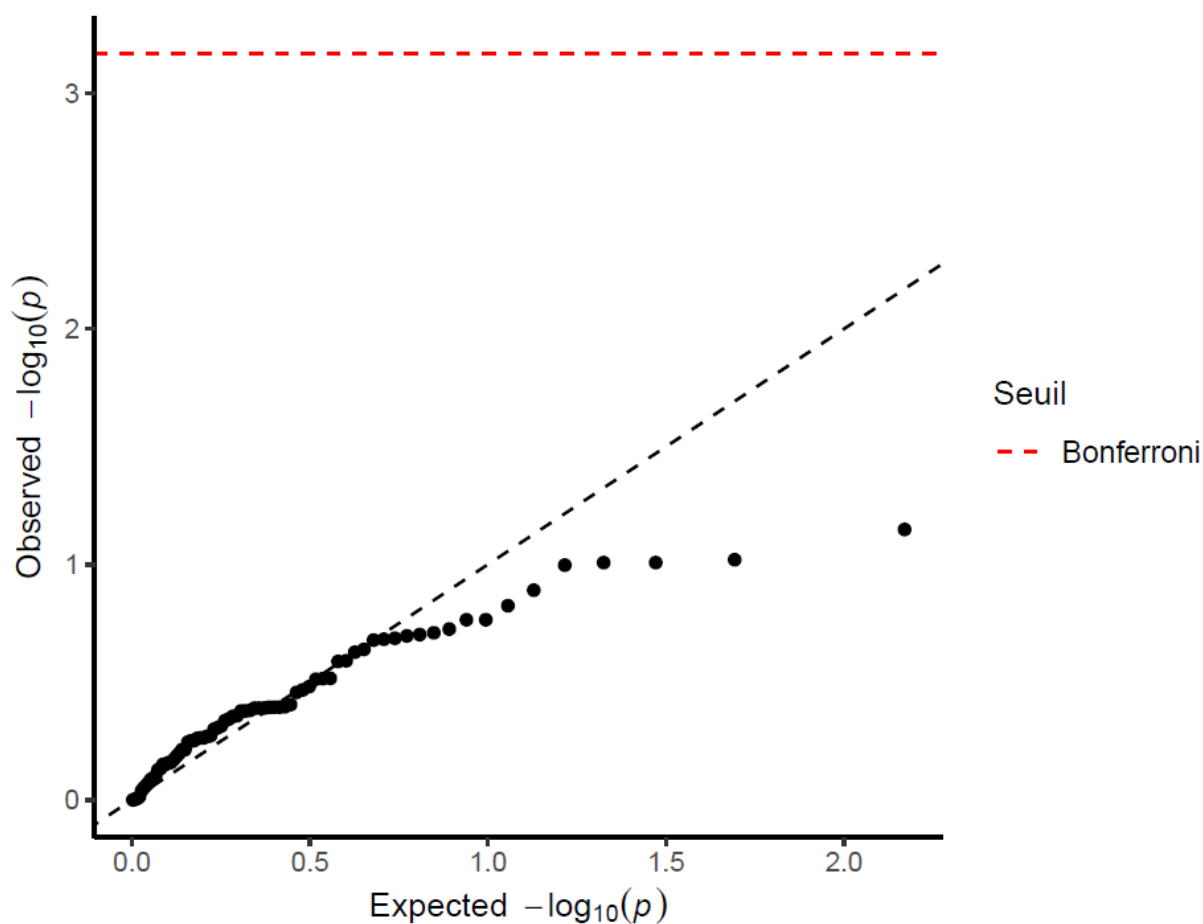


Figure III.D.3.19 : Volcano plot représentant les associations entre l'exposition à court terme au NO_2 et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

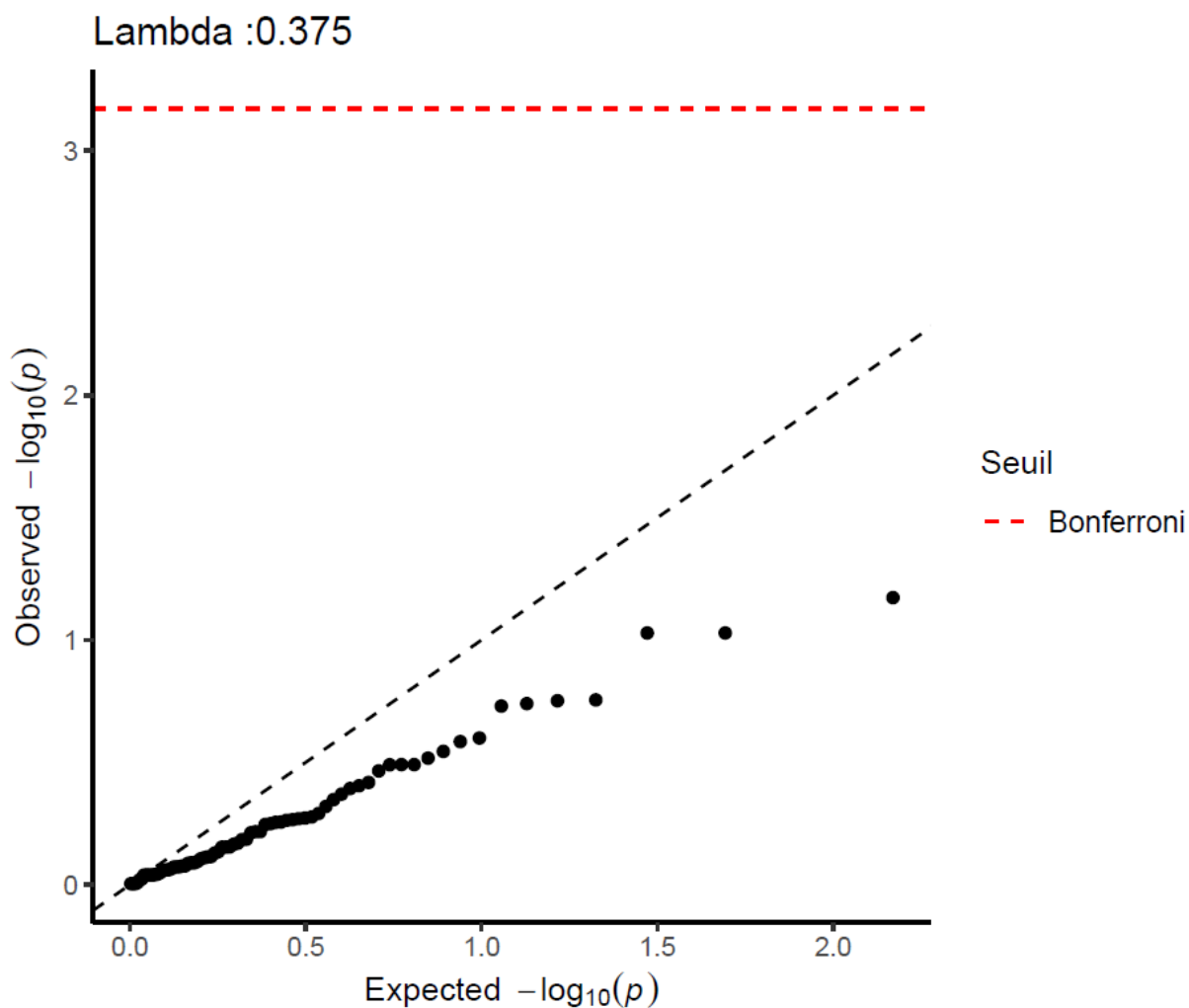


Figure III.D.3.20 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre l'exposition à court terme au NO₂ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

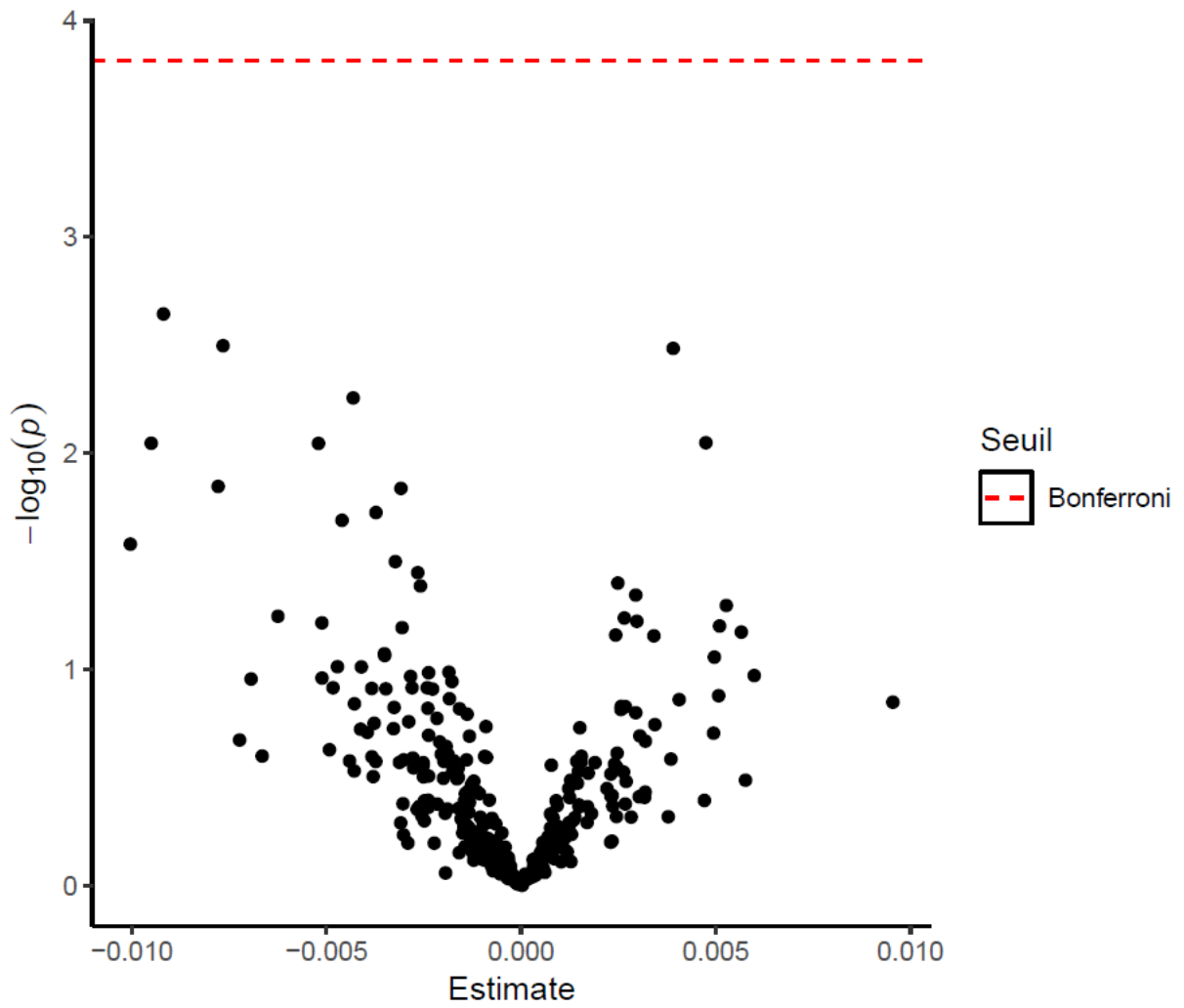


Figure III.D.3.21 : Volcano plot représentant les associations entre l'exposition résidentielle au NO₂ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

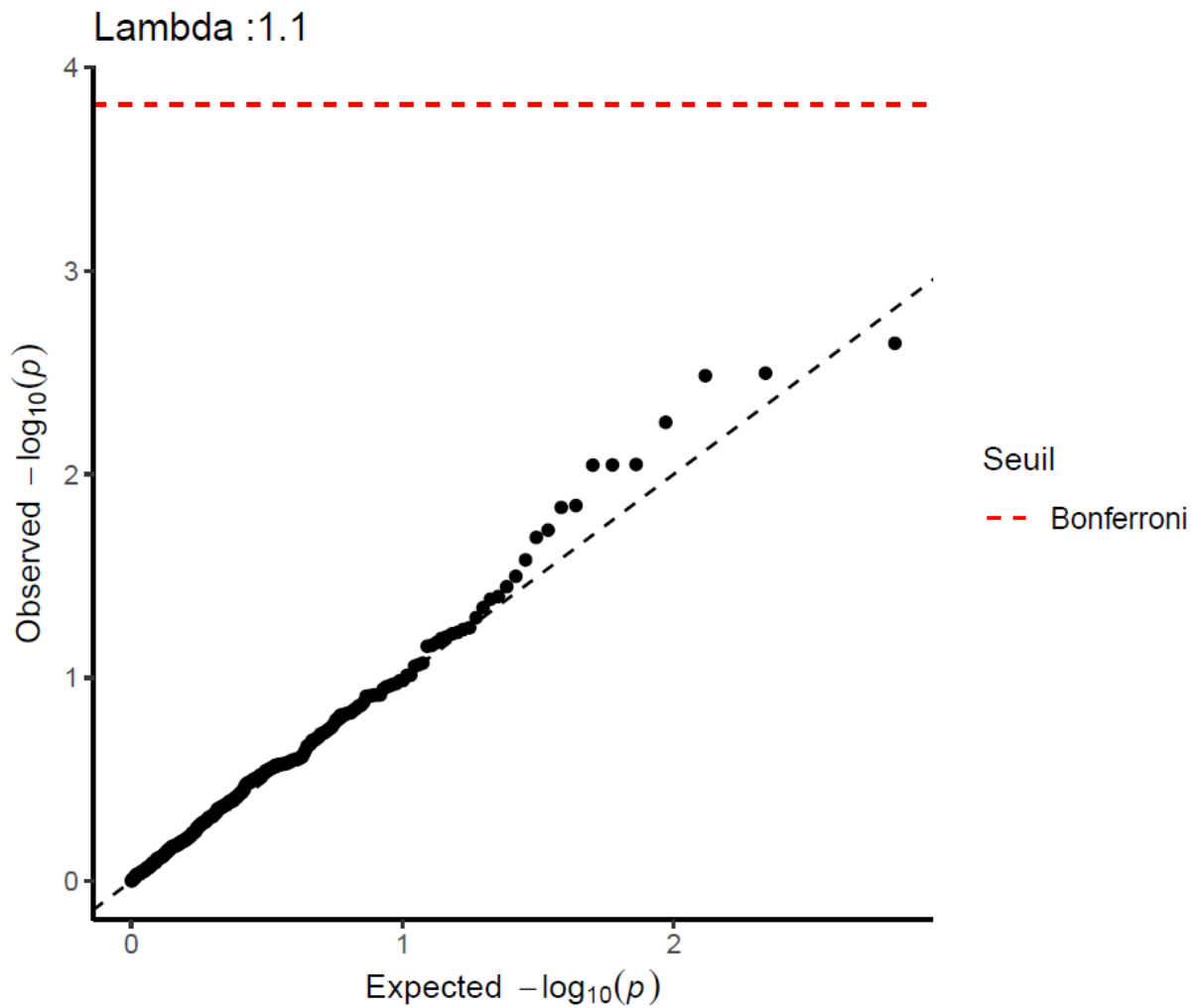


Figure III.D.3.22 : QQ plot représentant les associations entre l'exposition résidentielle au NO_2 et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

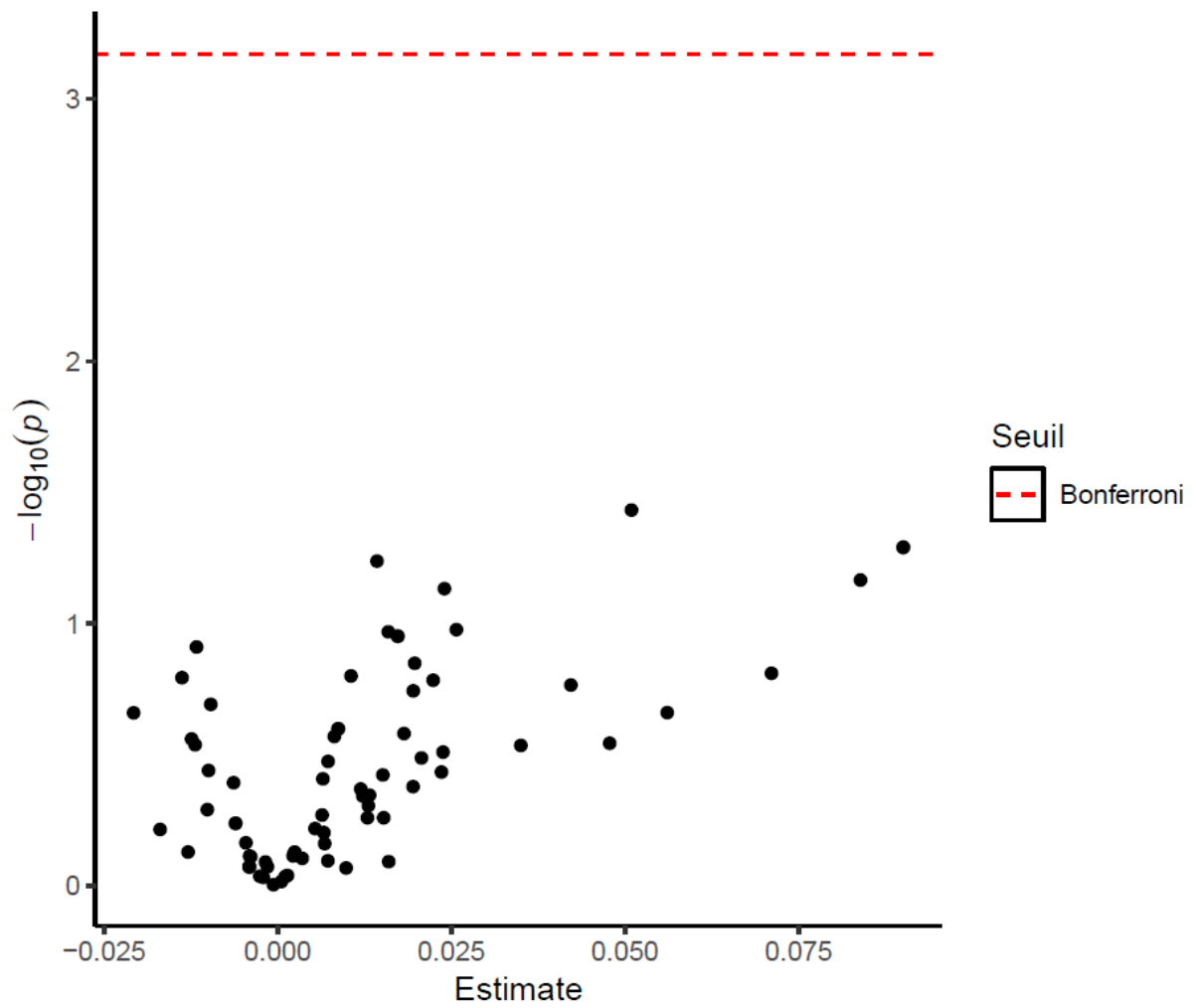


Figure III.D.3.23 : Volcano plot représentant les associations entre l'exposition résidentielle au NO₂ et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

Lambda : 1.48
Likelihood Ratio – pvalue : 0.781

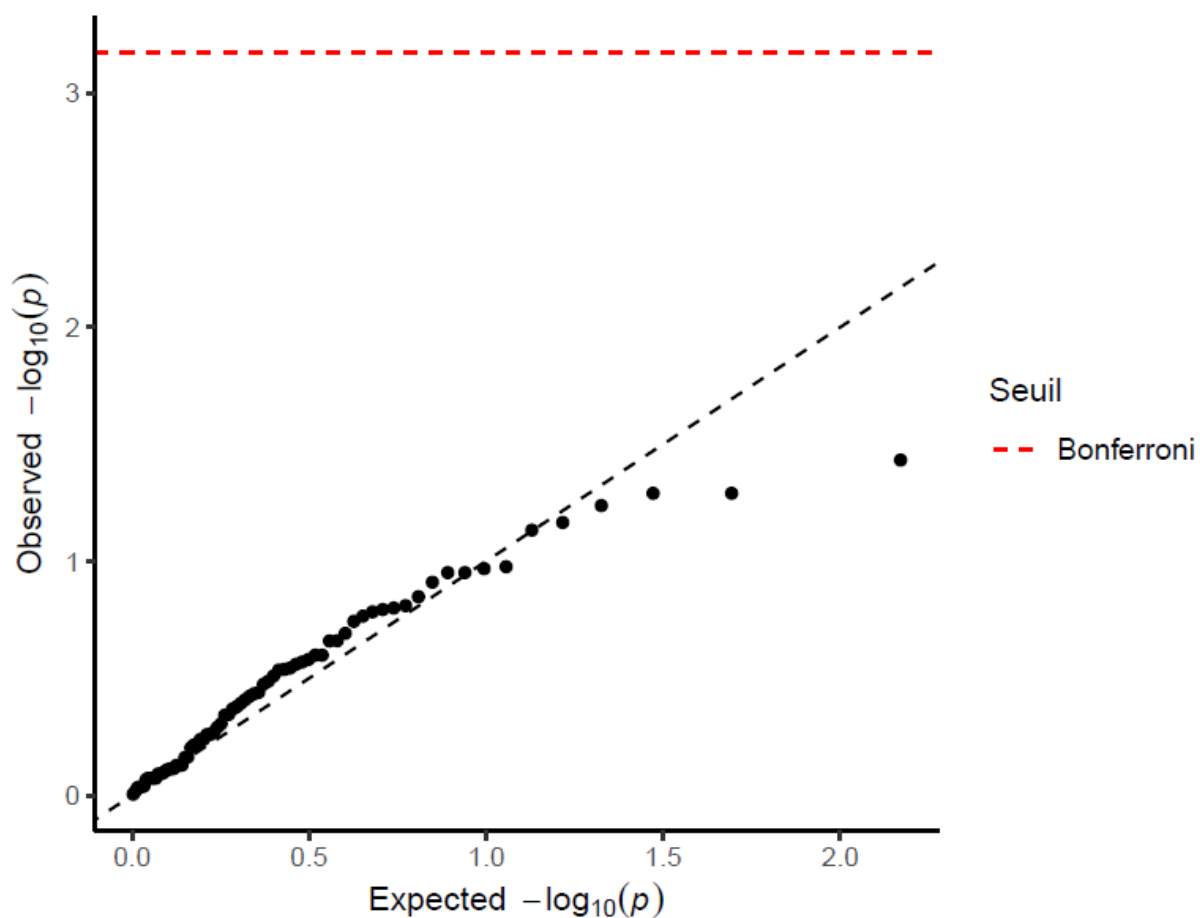


Figure III.D.3.24 : QQ plot représentant les associations entre l'exposition résidentielle au NO_2 et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

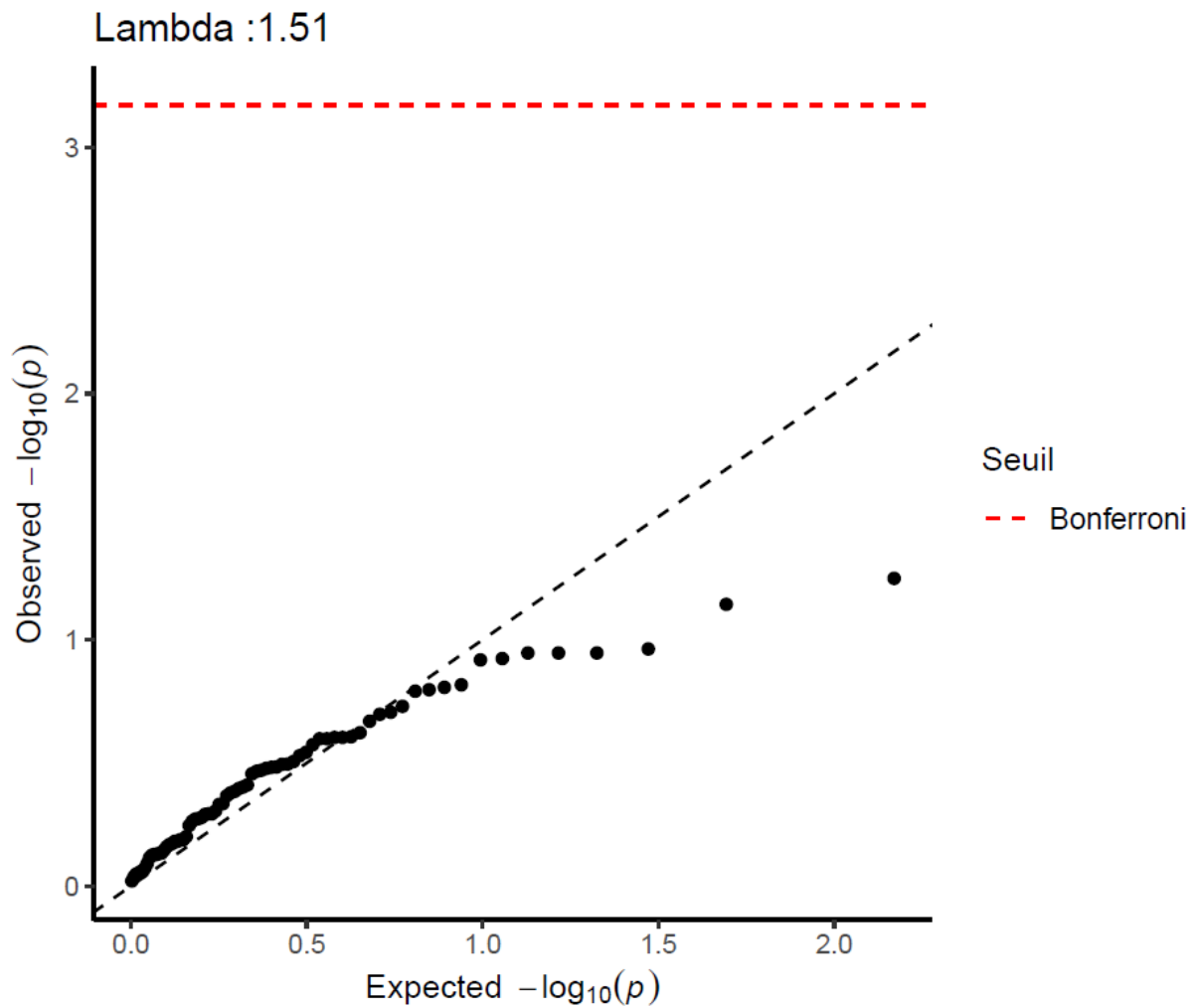


Figure III.D.3.25 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre l'exposition résidentielle au NO_2 et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation

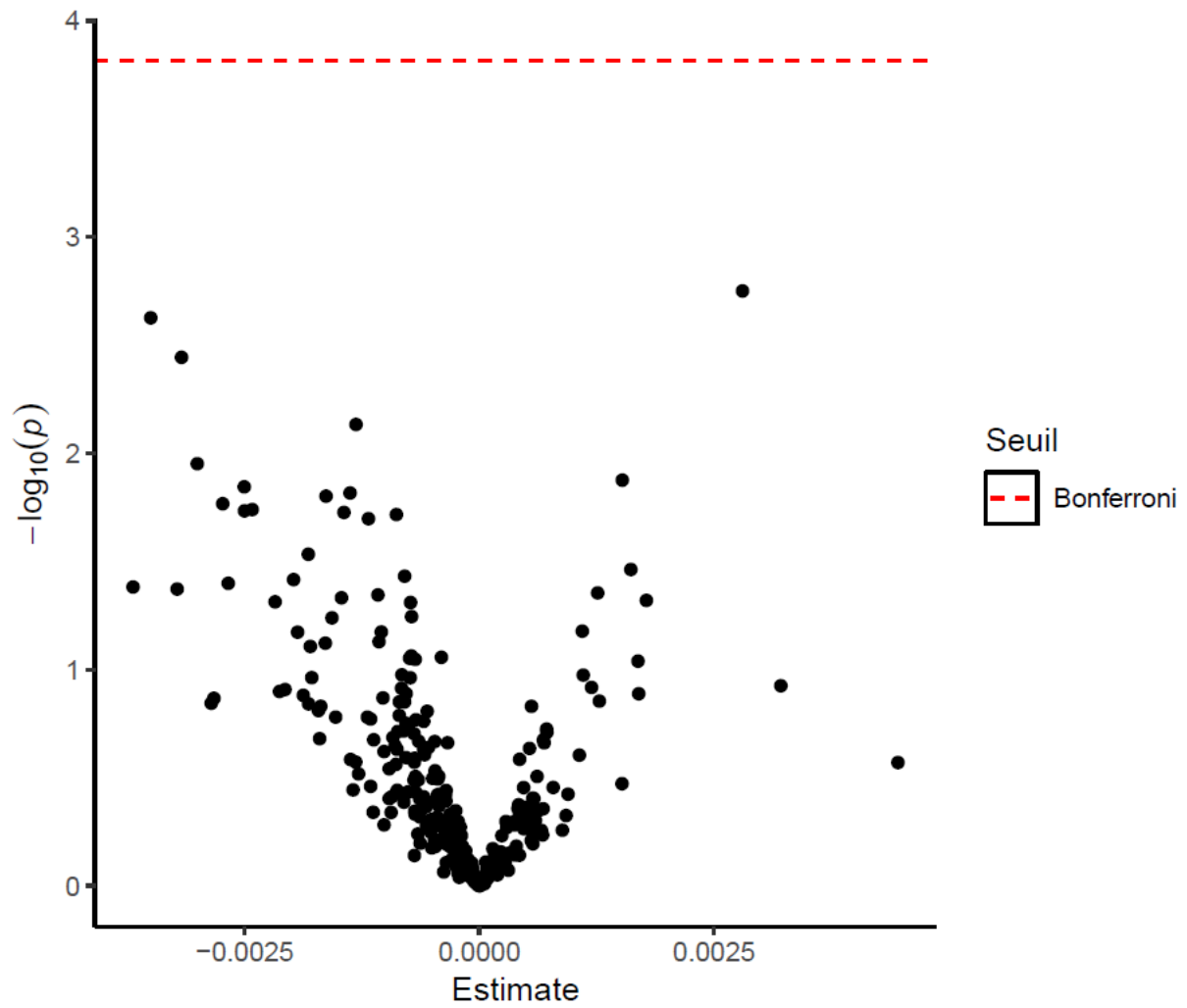


Figure III.D.3.26 : Volcano plot représentant les associations entre l'exposition à court terme à l'O₃ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

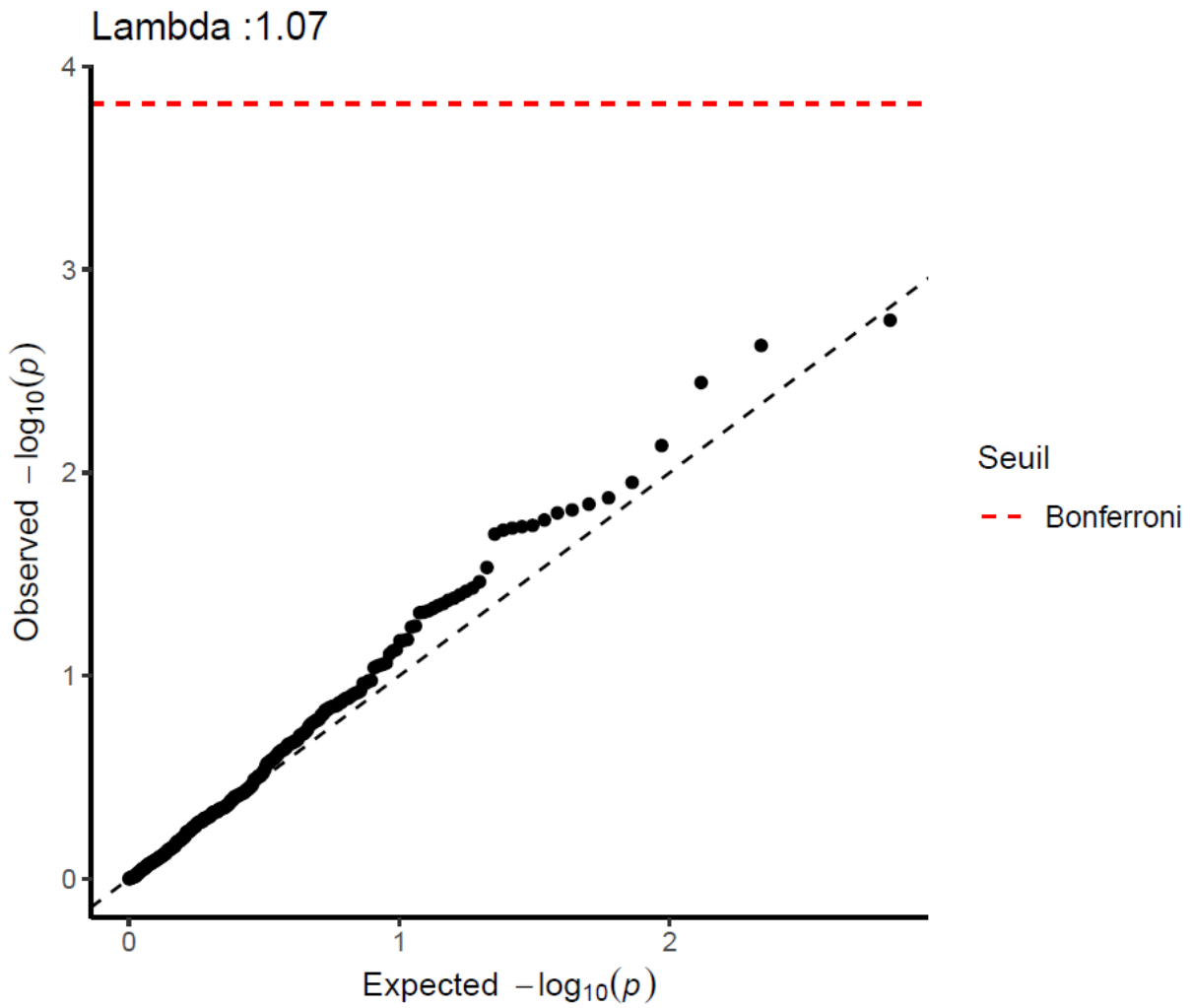


Figure III.D.3.27 : QQ plot représentant les associations entre l'exposition à court terme à l'O₃ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

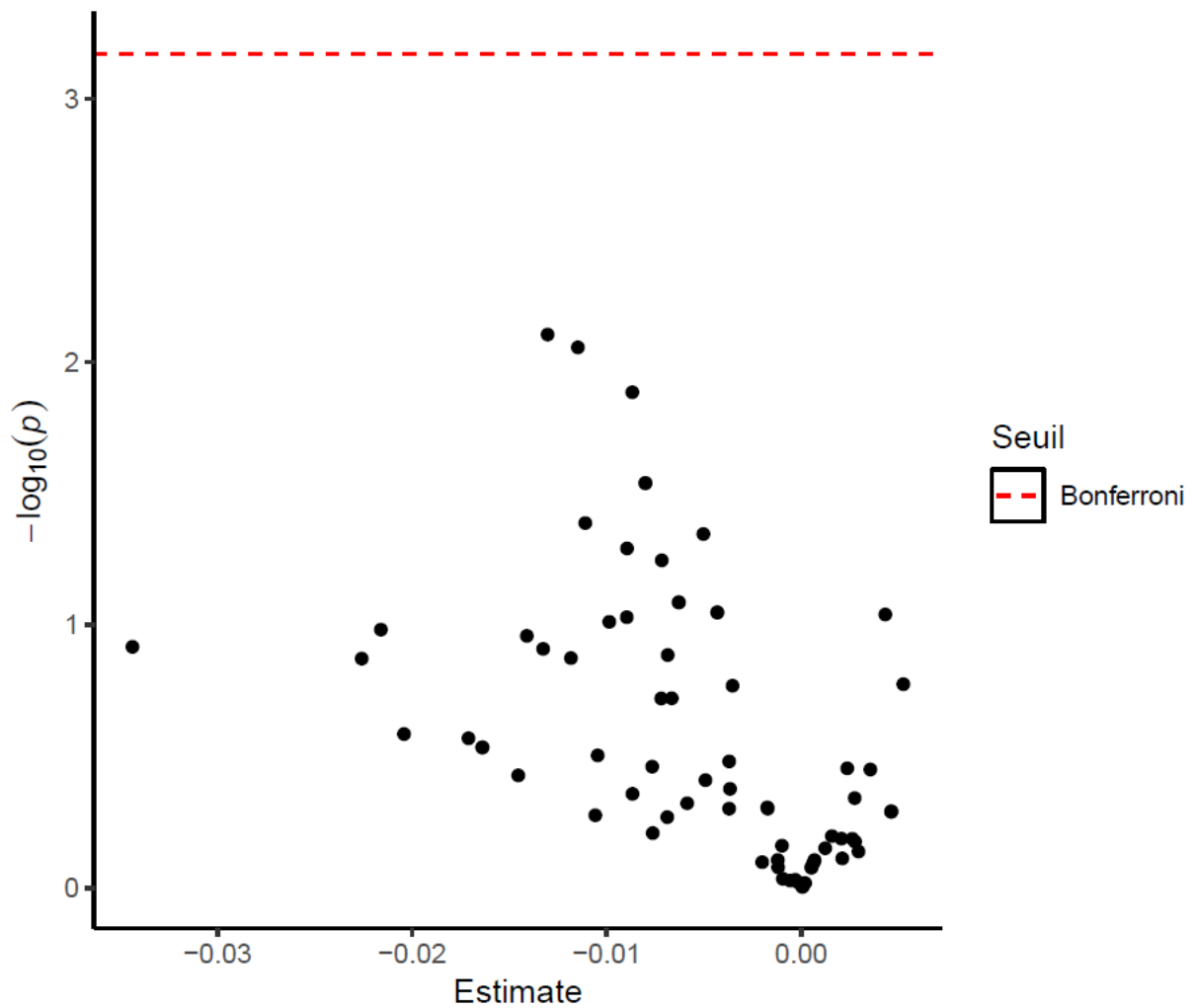


Figure III.D.3.28 : Volcano plot représentant les associations entre l'exposition à court terme à l'O₃ et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

Lambda : 1.69
Likelihood Ratio – pvalue : 0.89

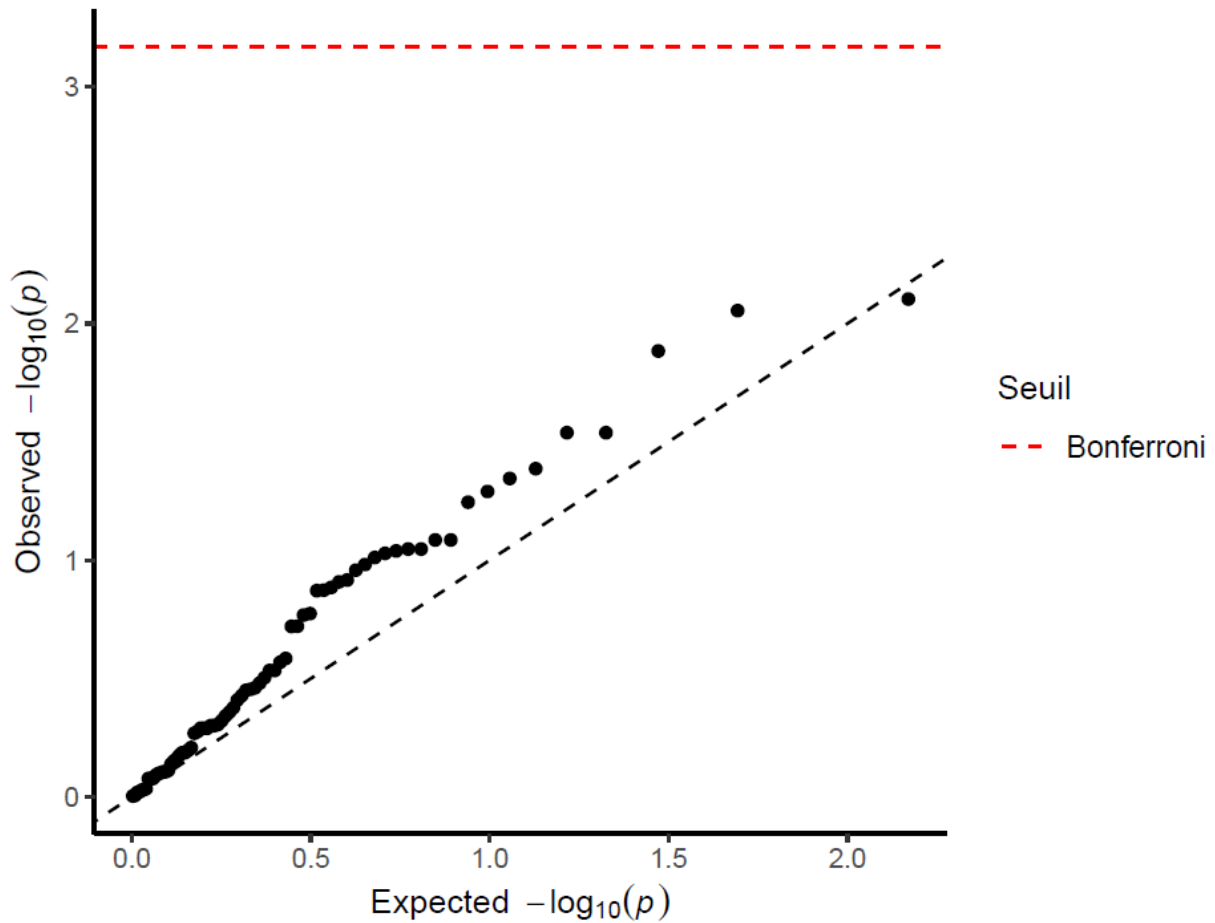


Figure III.D.3.29 : QQ plot représentant les associations entre l'exposition à court terme à l'O₃ et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

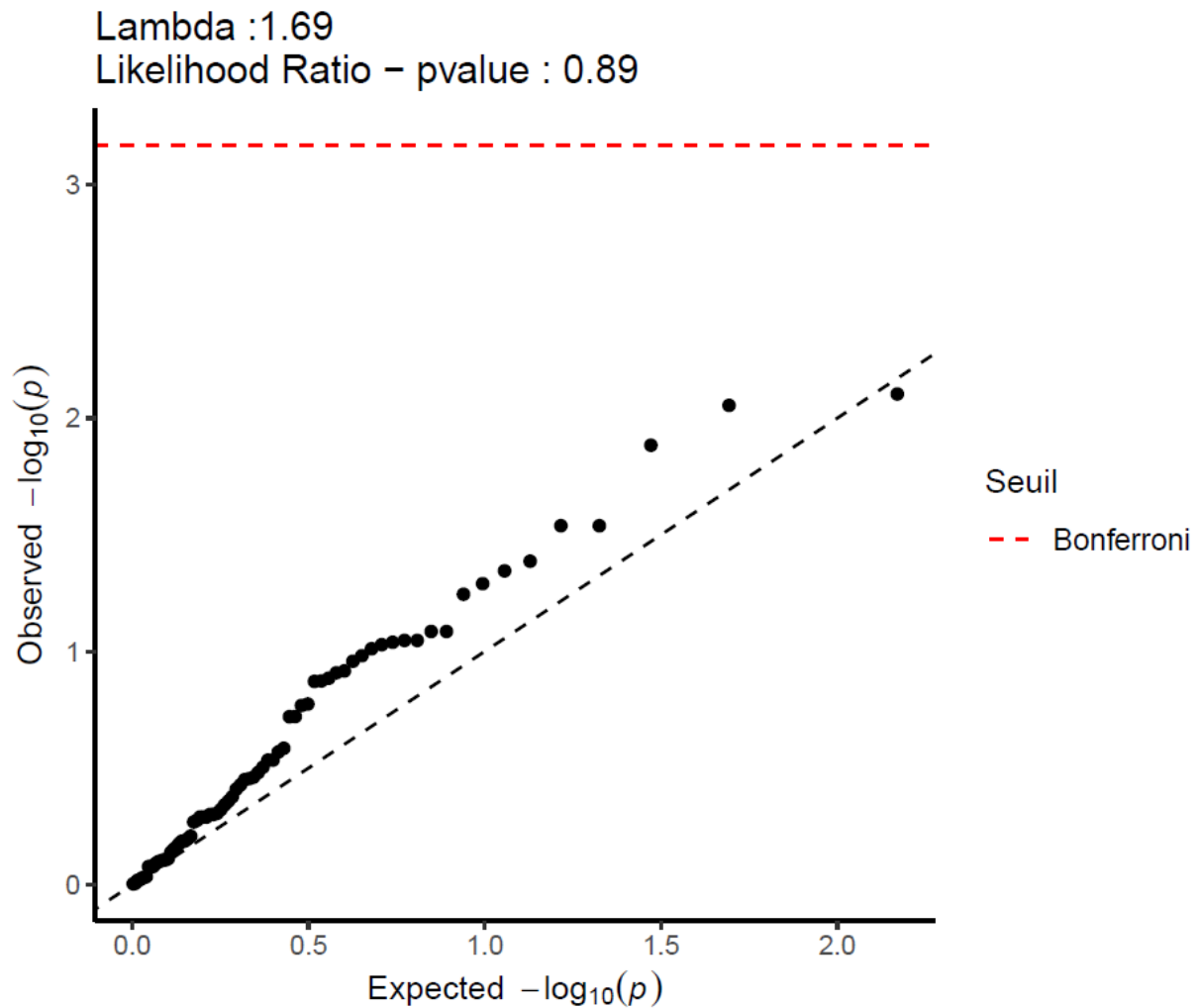


Figure III.D.3.30 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre l'exposition à court terme à l'O₃ et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation

(3) Exposition domestique

Nous avons regardé les associations entre la présence de moisissures dans le logement des participants. Aucune association n'était significative avec les métabolites ou les voies métaboliques (Figures III.D.3.31, III.D.3.32, III.D.3.33, III.D.3.34 ; III.D.3.35).

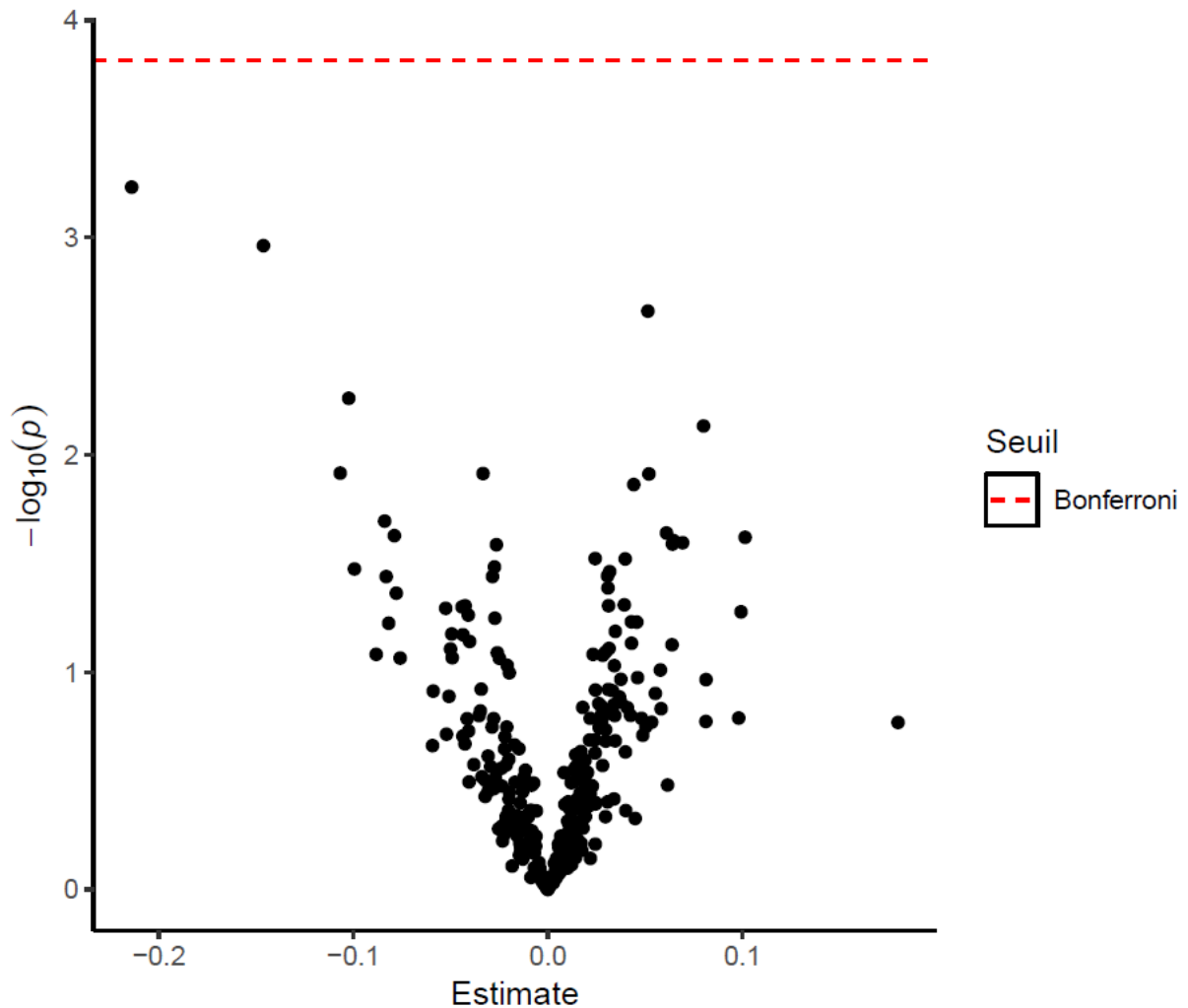


Figure III.D.3.31 : Volcano plot représentant les associations entre les moisissures et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

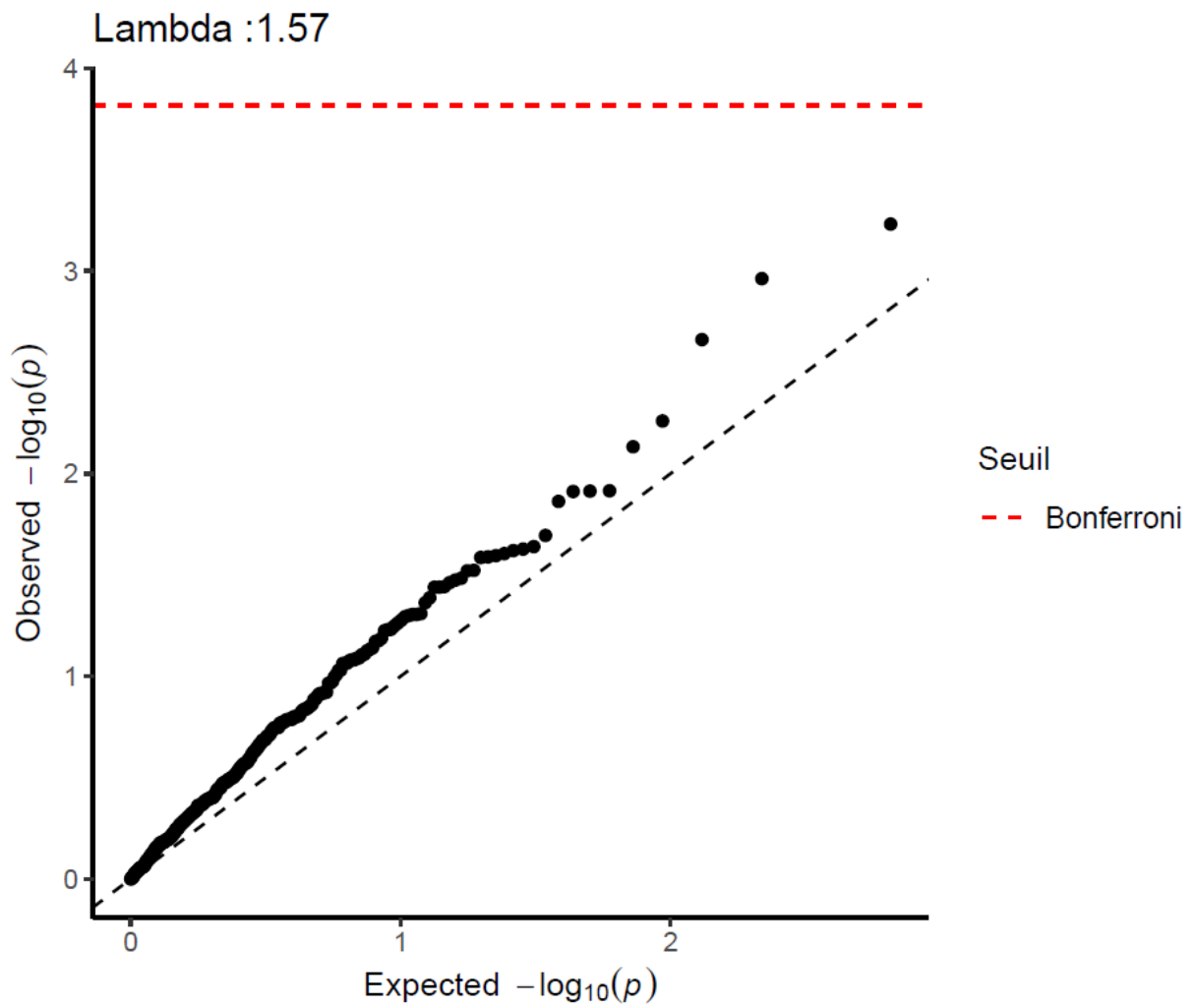


Figure III.D.3.32 : QQ plot représentant les associations entre les moisissures et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

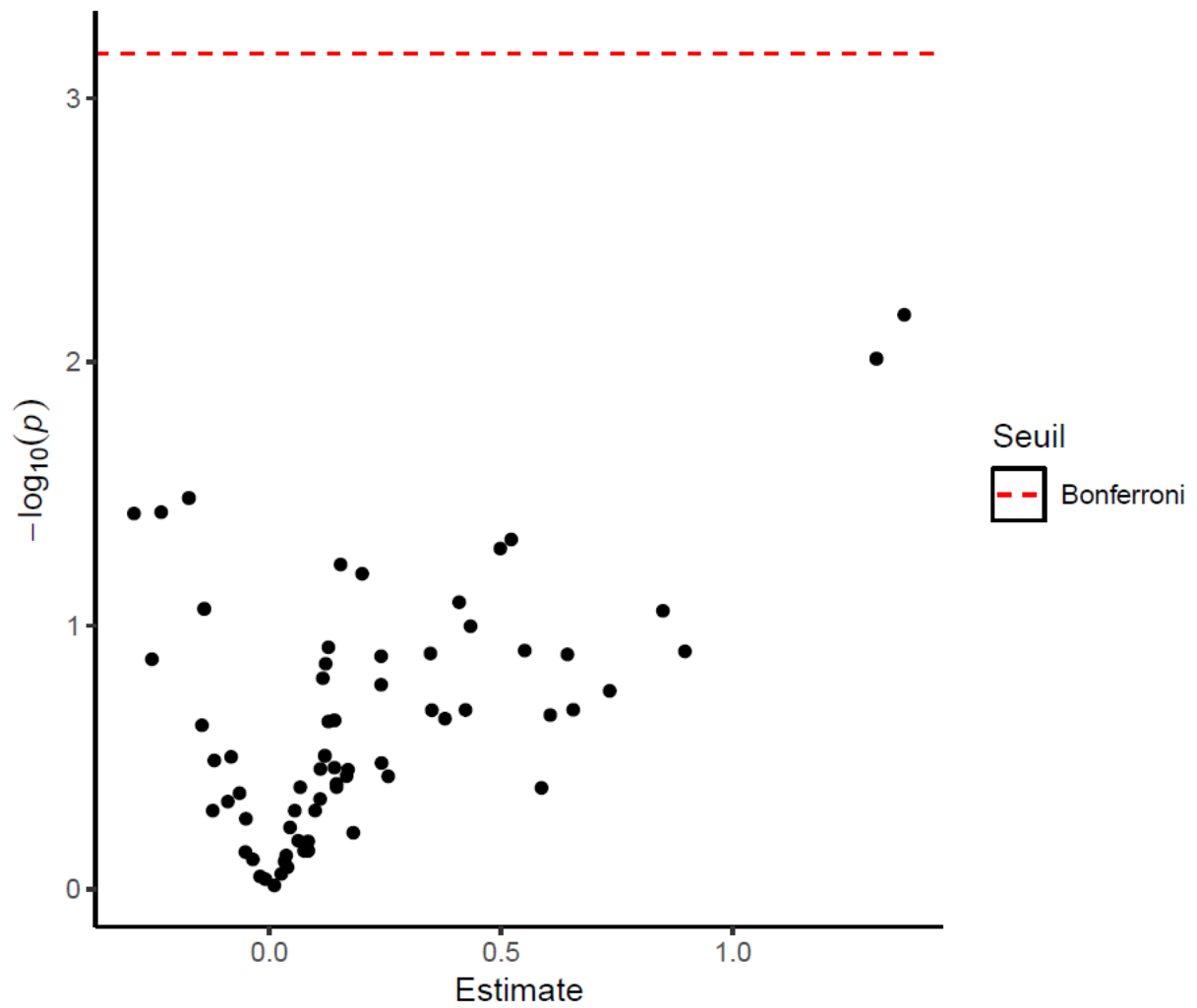


Figure III.D.3.33 : Volcano plot représentant les associations entre les moisissures et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

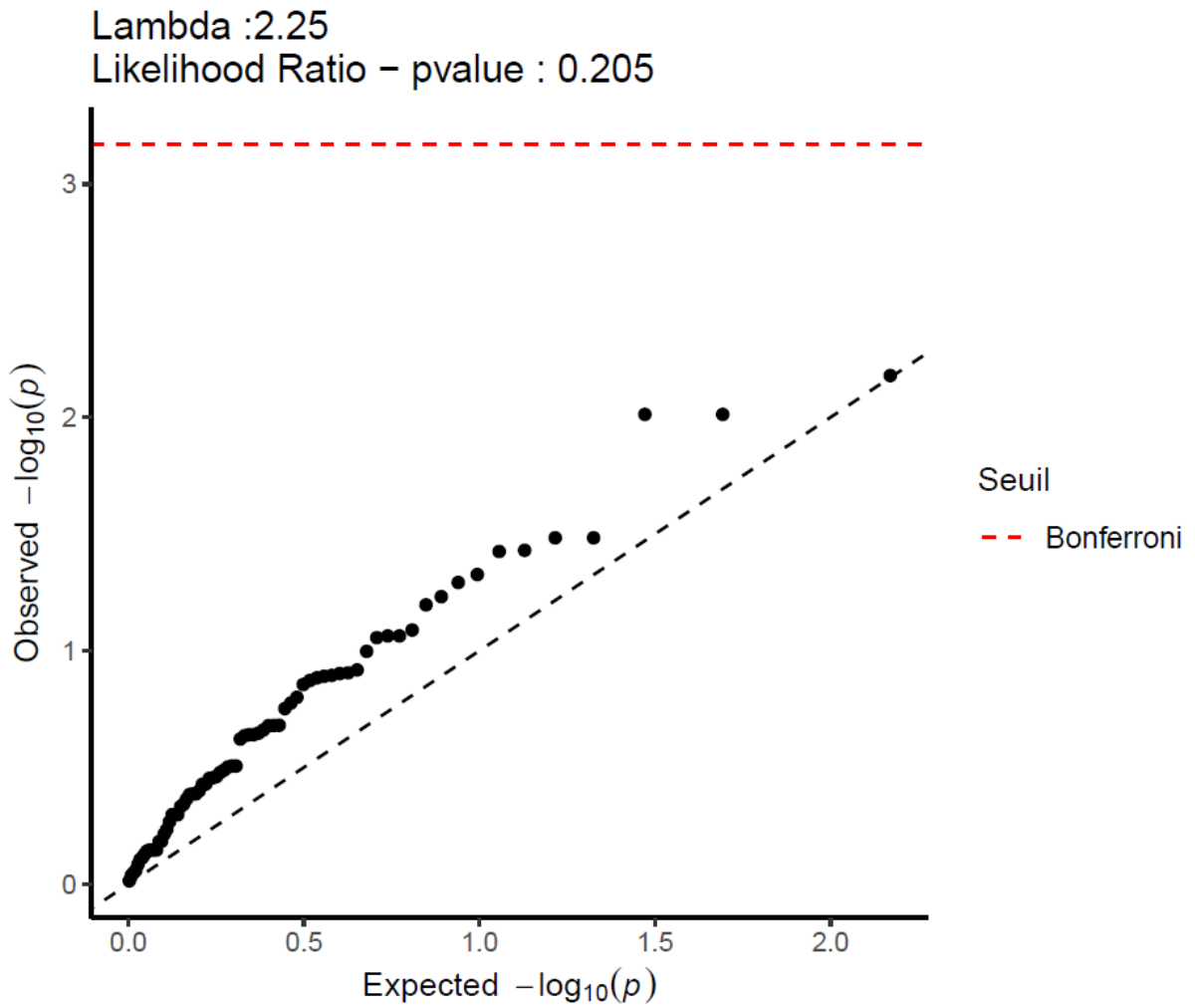


Figure III.D.3.34 : QQ plot représentant les associations entre les moisissures et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

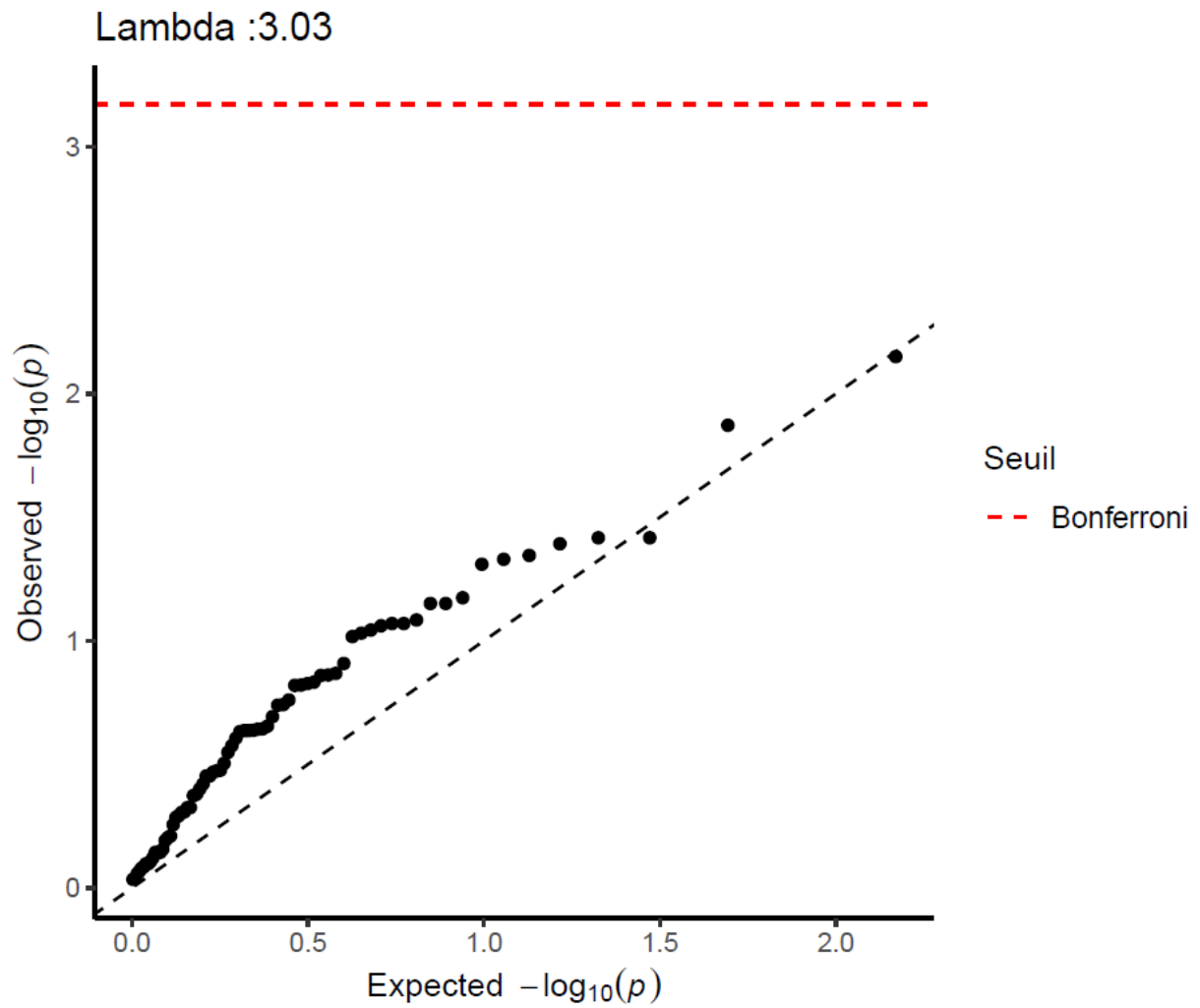


Figure III.D.3.35 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre les moisissures et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine

(4) Exposition professionnelle

Nous avons retrouvé des associations significatives entre l'exposition aux vapeurs et 3 métabolites ou classe de métabolites : la classe des hydroxyindoles, la classe des acides méthyls hippuriques, et le sulfate de glucuronide d'estradiol (figures III.D.3.36 et III.D.3.37). La classe des acides méthyls hippuriques est augmentée avec l'exposition aux vapeurs, tandis que la classe des hydroxyindoles et le sulfate de glucuronide d'estradiol sont diminués.

Nous n'avons pas de précision sur le contenu des vapeurs. Il est donc difficile de tirer des conclusions sur les métabolites associés à cette exposition. Il existe un lien connu entre les acides méthyls hippuriques et l'exposition au xylène, un solvant utilisé dans l'industrie (260). Pour les autres métabolites retrouvés, il est difficile de formuler des hypothèses à ce stade. Les associations entre l'exposition aux vapeurs et les voies métaboliques n'étaient pas significatives (figures III.D.3.38, III.D.3.39 et III.D.3.40)

L'exposition aux poussières n'était pas associée aux métabolites (figures III.D.3.41 et III.D.3.42). Cependant, elle était associée à 3 voies métaboliques (figures III.D.3.41, III.D.3.42) : la voie du métabolisme des eicosanoïdes, la voie du métabolisme de la thiamine et la voie du métabolisme du glutathion. Ces voies métaboliques jouent un rôle dans la régulation du stress oxydant (261–263). La représentation des p des rapports de vraisemblance montre une déviation suggérant de nombreuses associations entre les voies métaboliques et l'exposition aux poussières (figure III.D.3.45).

A l'inverse, les associations étaient non significatives avec les expositions professionnelles aux gaz (figures III.D.3.46, III.D.3.47, III.D.3.48 et III.D.3.50) et aux fumées (figures III.D.3.51, III.D.3.52, III.D.3.53, III.D.3.54 et III.D.3.55). Cependant, le rapport de vraisemblance sur l'ensemble des voies métaboliques et l'exposition aux fumées était significatif ($p = 0,0342$) (figure III.D.54). Ce résultat suggère que l'exposition aux fumées impacte au moins une voie métabolique.

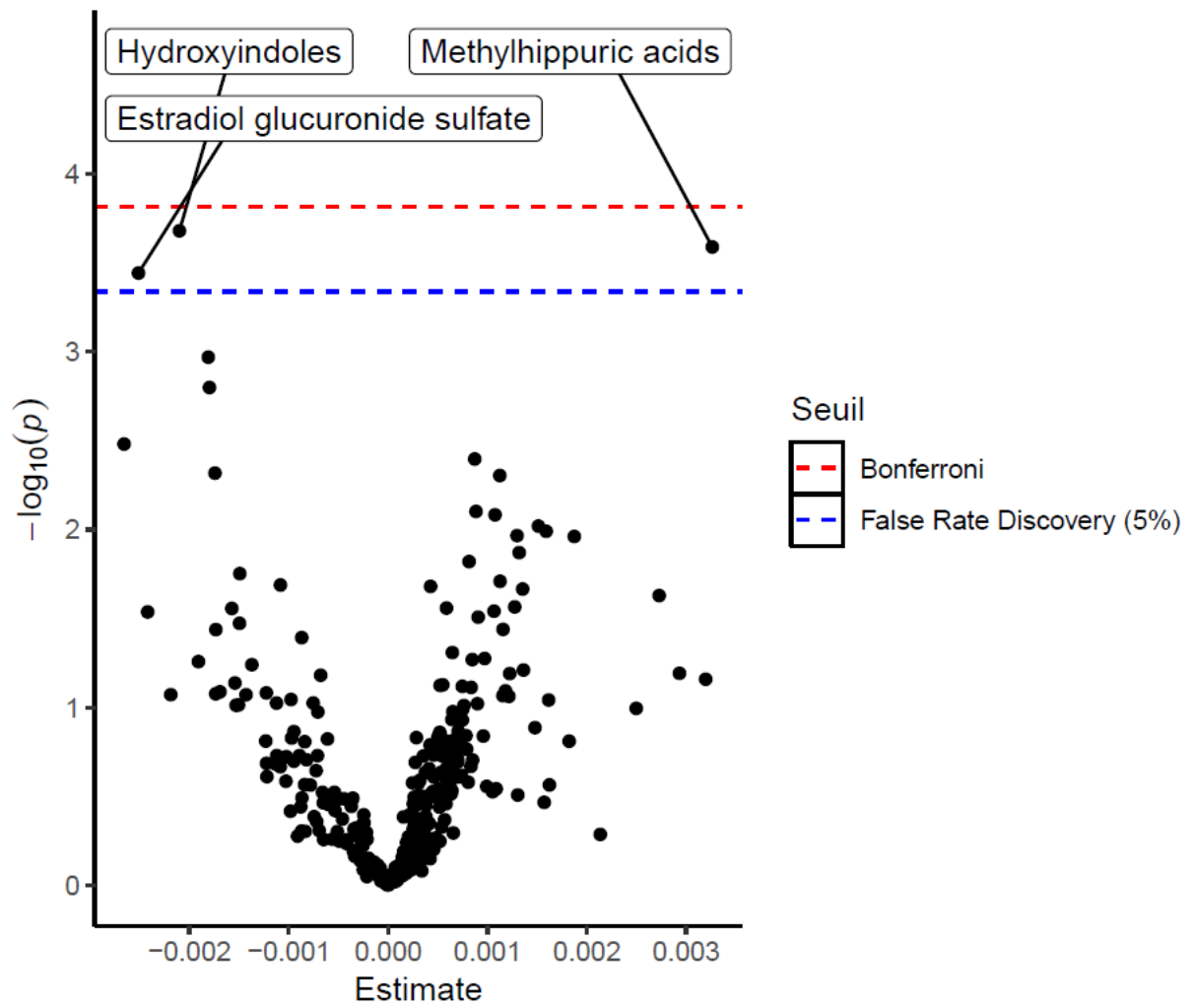


Figure III.D.3.36 : Volcano plot représentant les associations entre les vapeurs et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

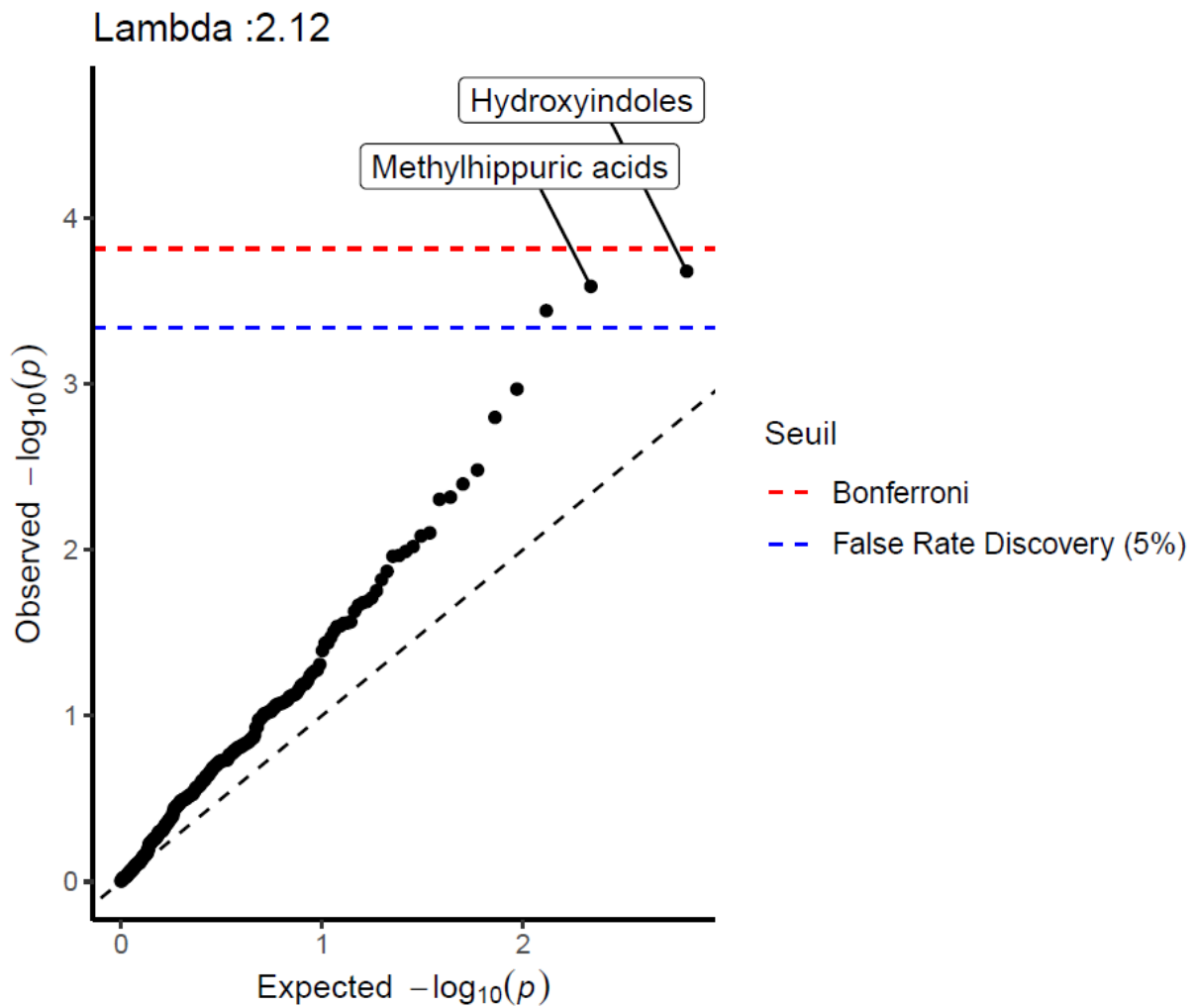


Figure III.D.3.37 : QQ plot représentant les associations entre les vapeurs et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

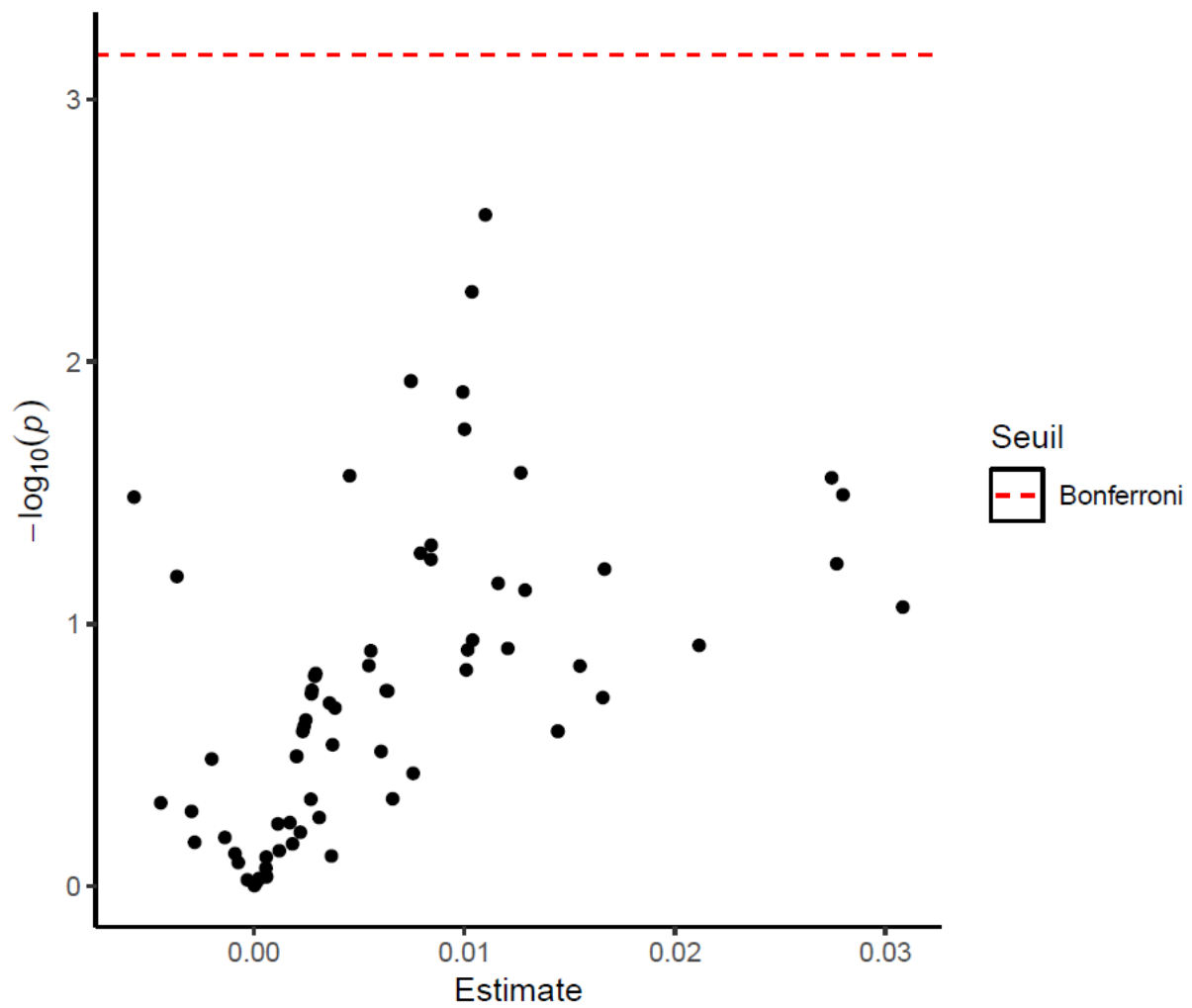


Figure III.D.3.38 : Volcano plot représentant les associations entre les vapeurs et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

Lambda :3.54
Likelihood Ratio – pvalue : 0.14

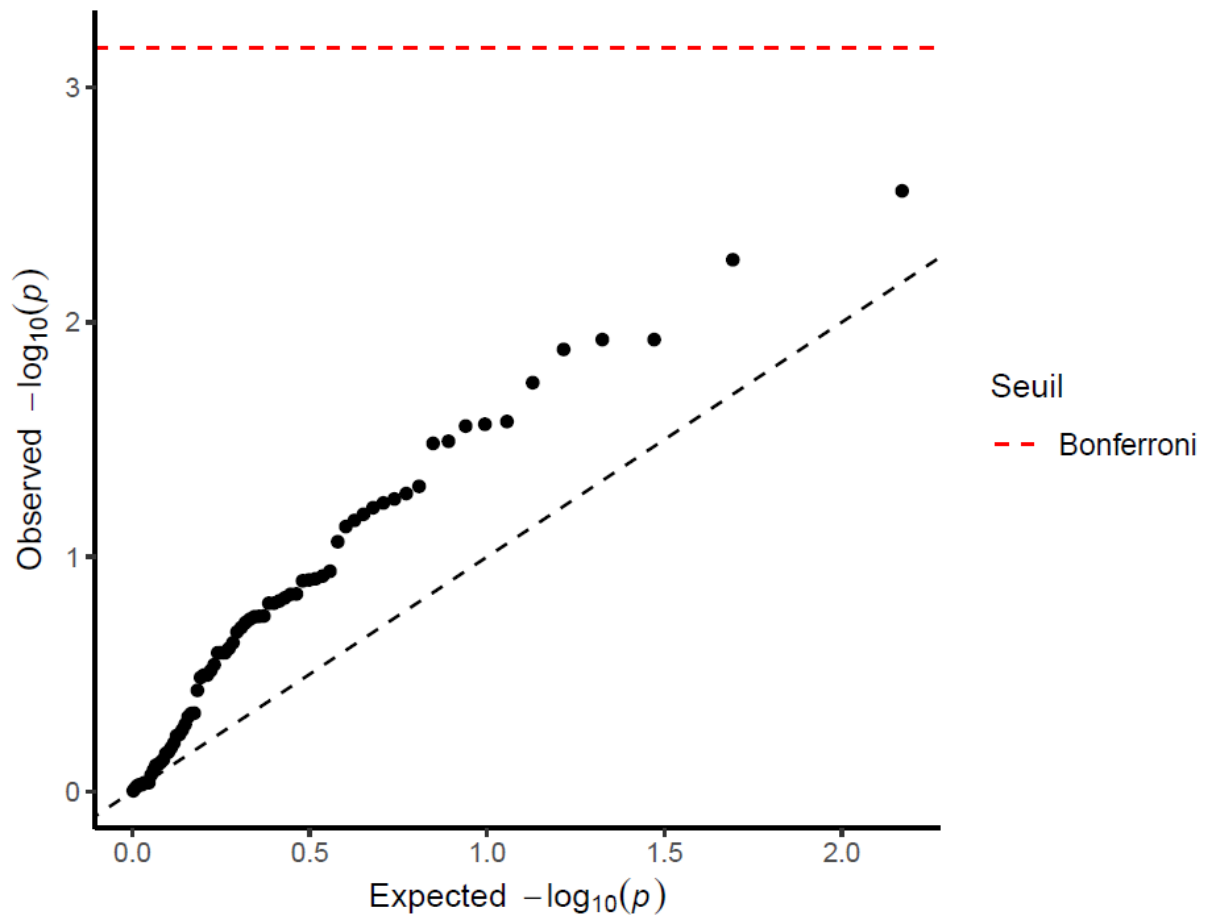


Figure III.D.3.39 : QQ plot représentant les associations entre les vapeurs et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

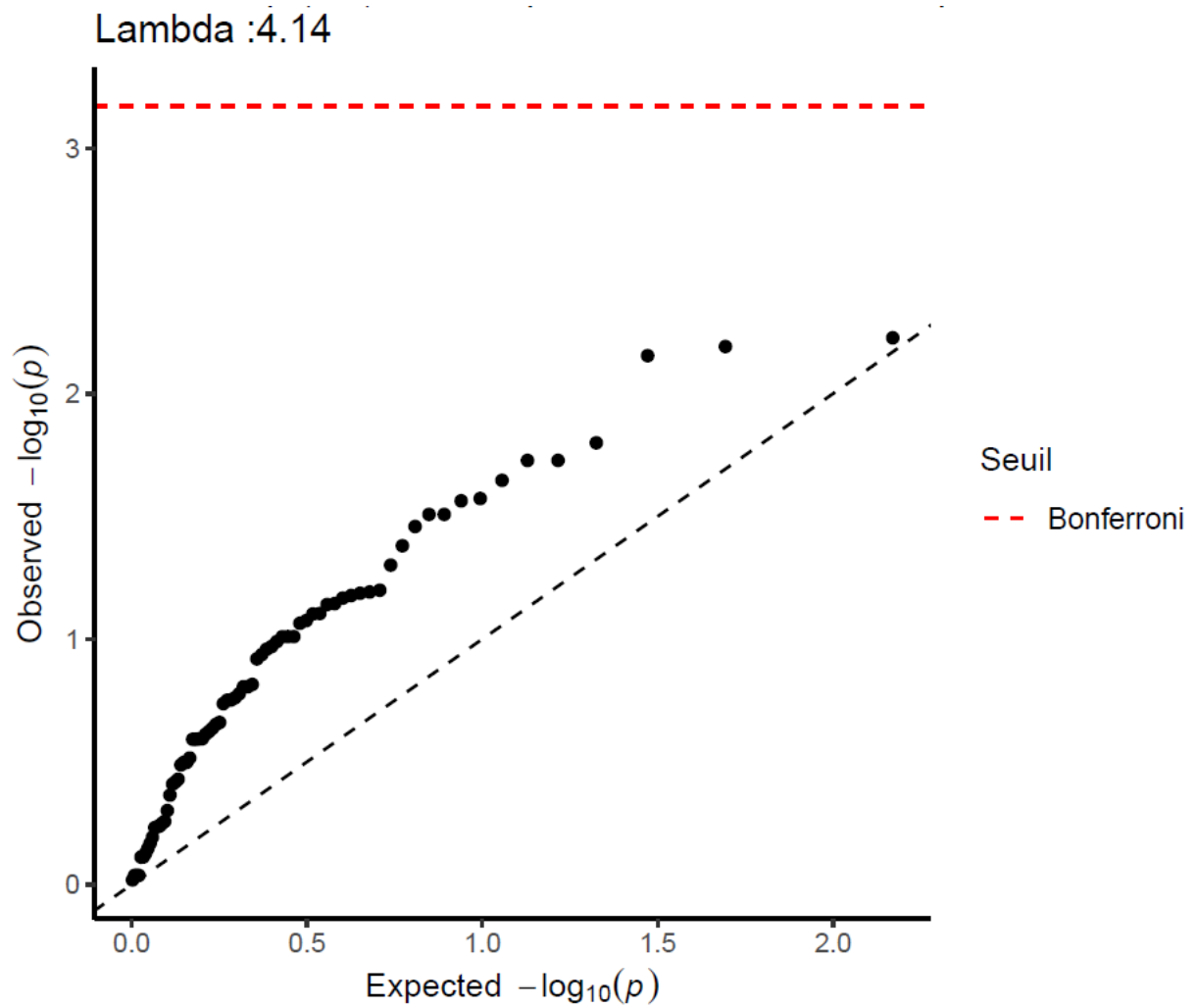


Figure III.D.3.40 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre les vapeurs et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine

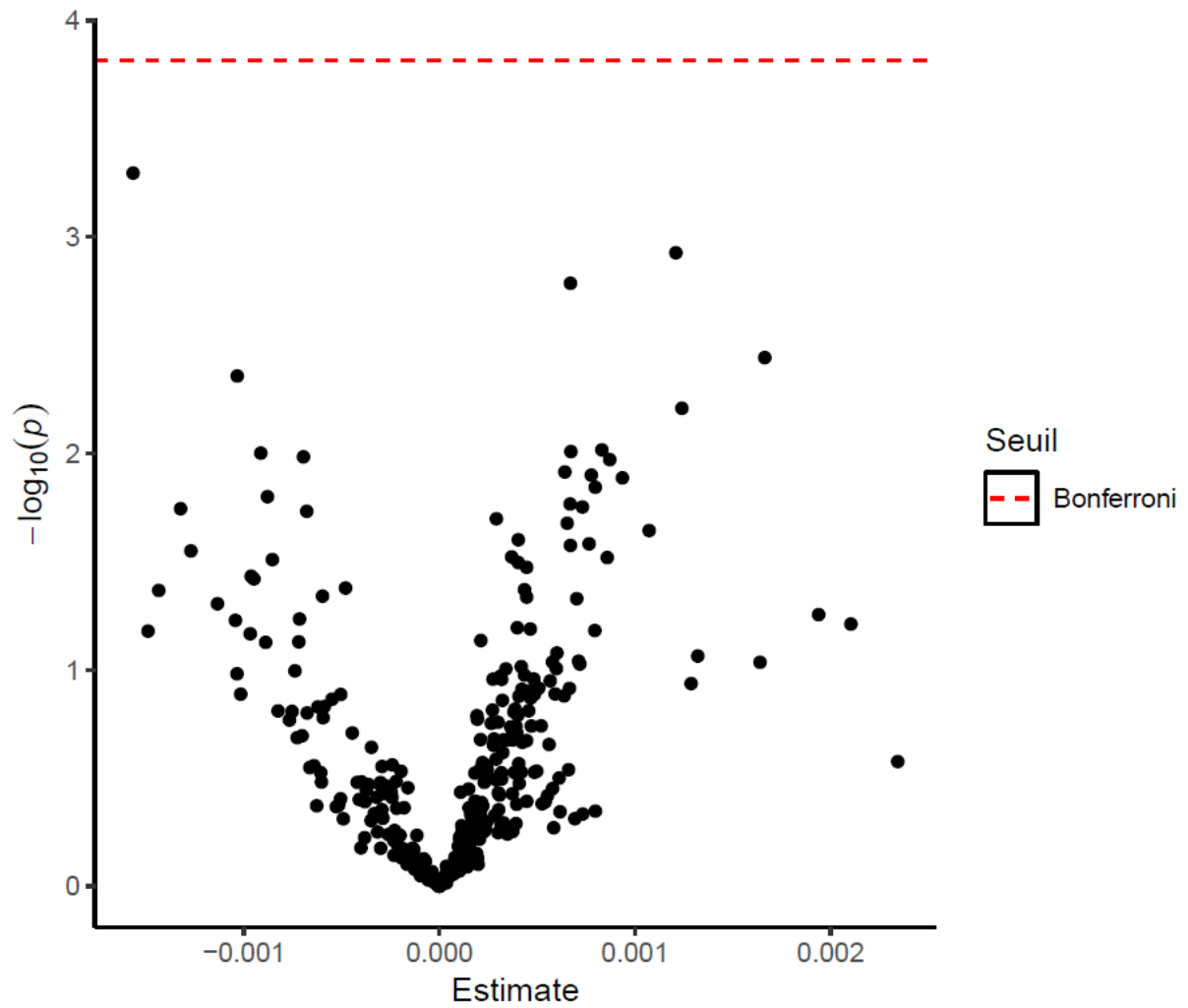


Figure III.D.3.41 : Volcano plot représentant les associations entre les poussières et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

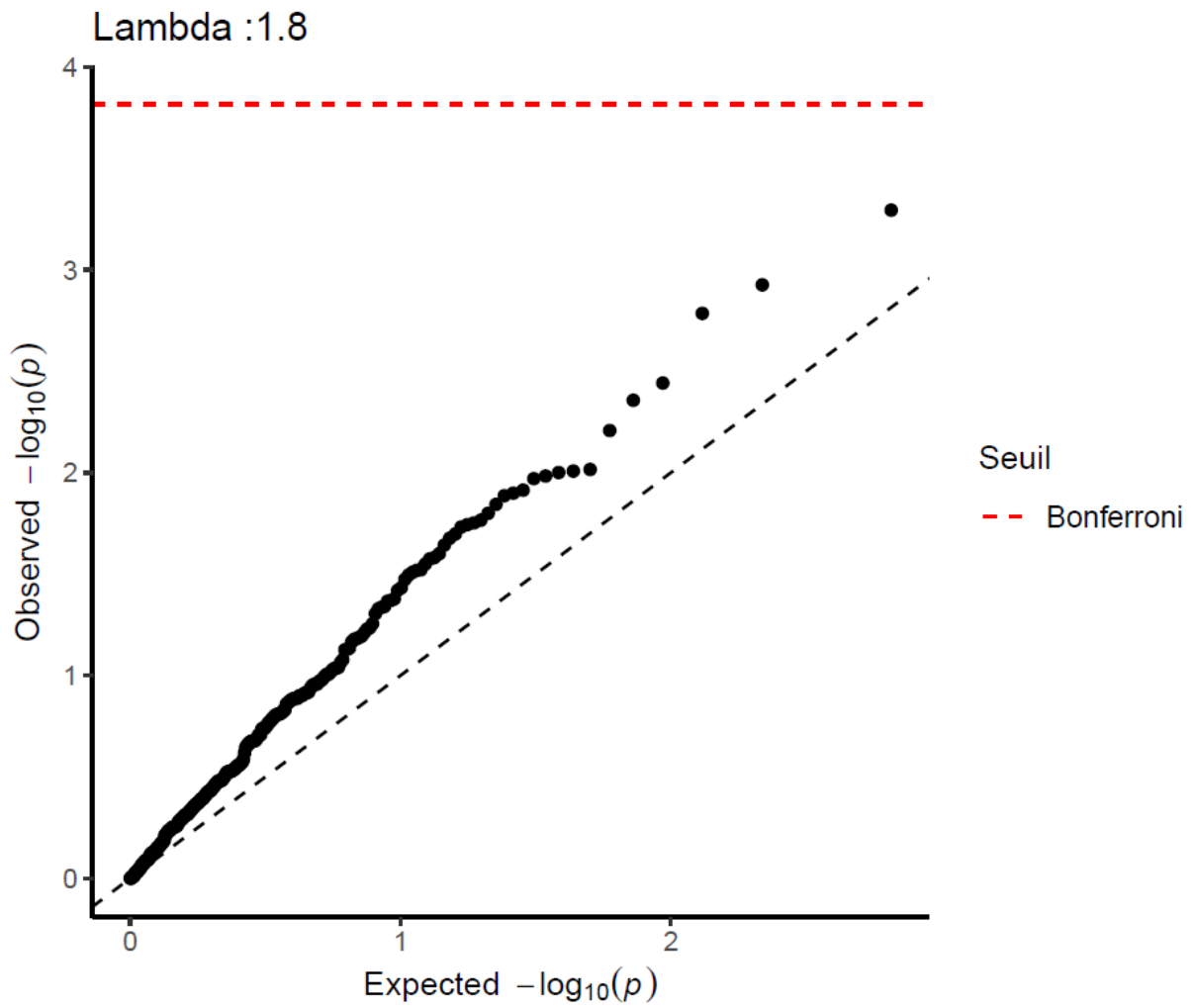


Figure III.D.3.42 : QQ plot représentant les associations entre les poussières et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

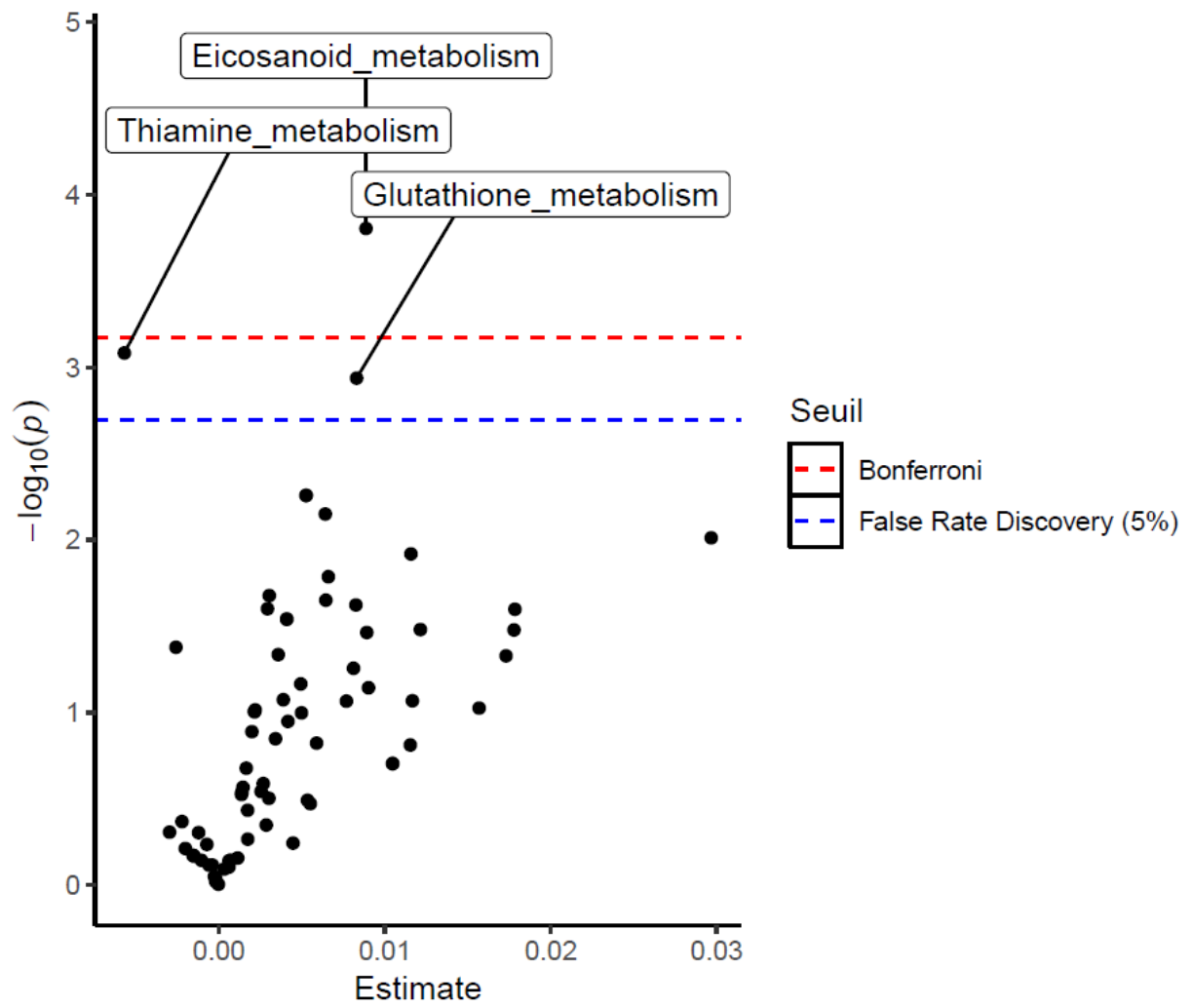


Figure III.D.3.43 : Volcano plot représentant les associations entre les poussières et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

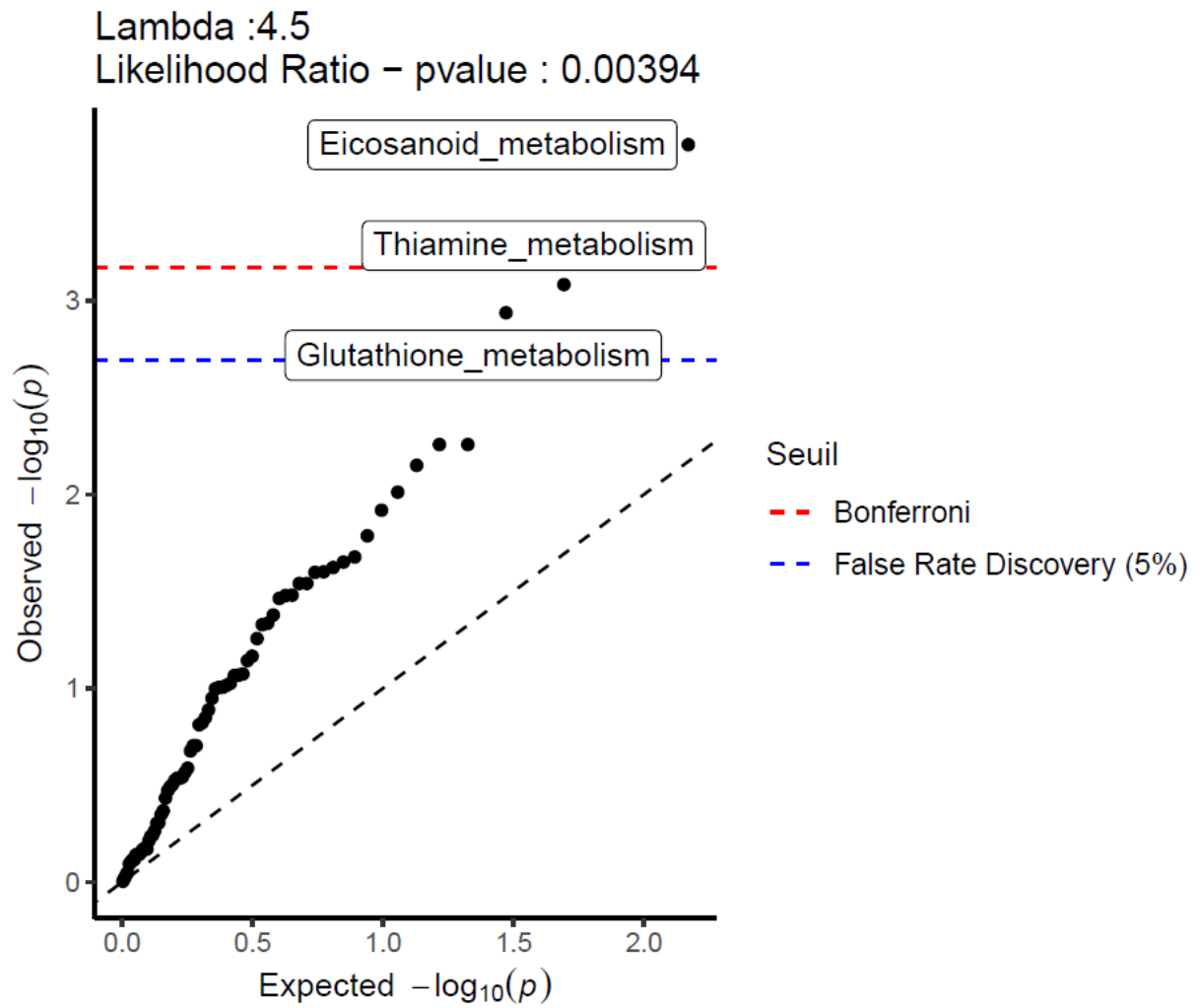


Figure III.D.3.44 : QQ plot représentant les associations entre les poussières et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

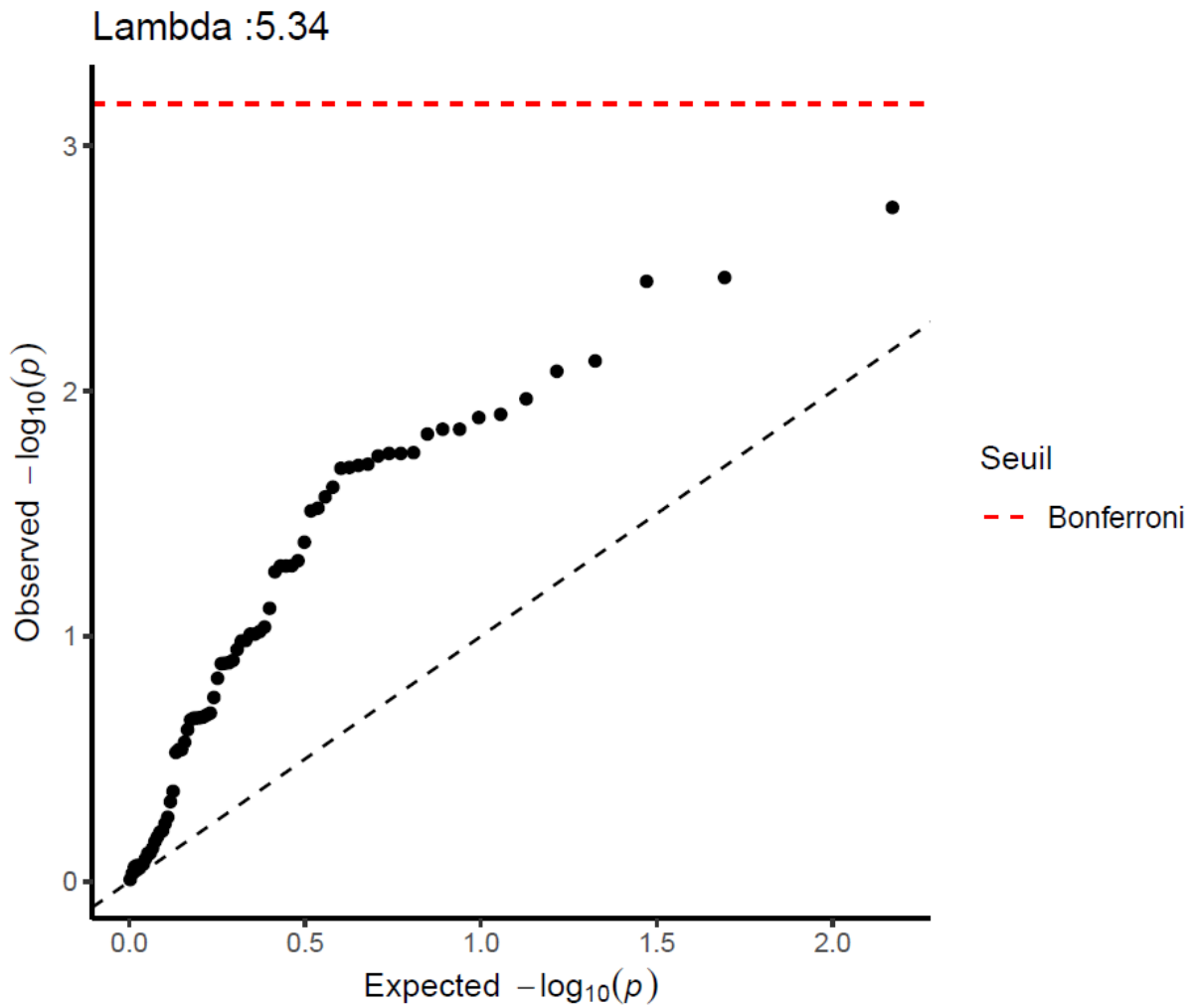


Figure III.D.3.45 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre les poussières et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine

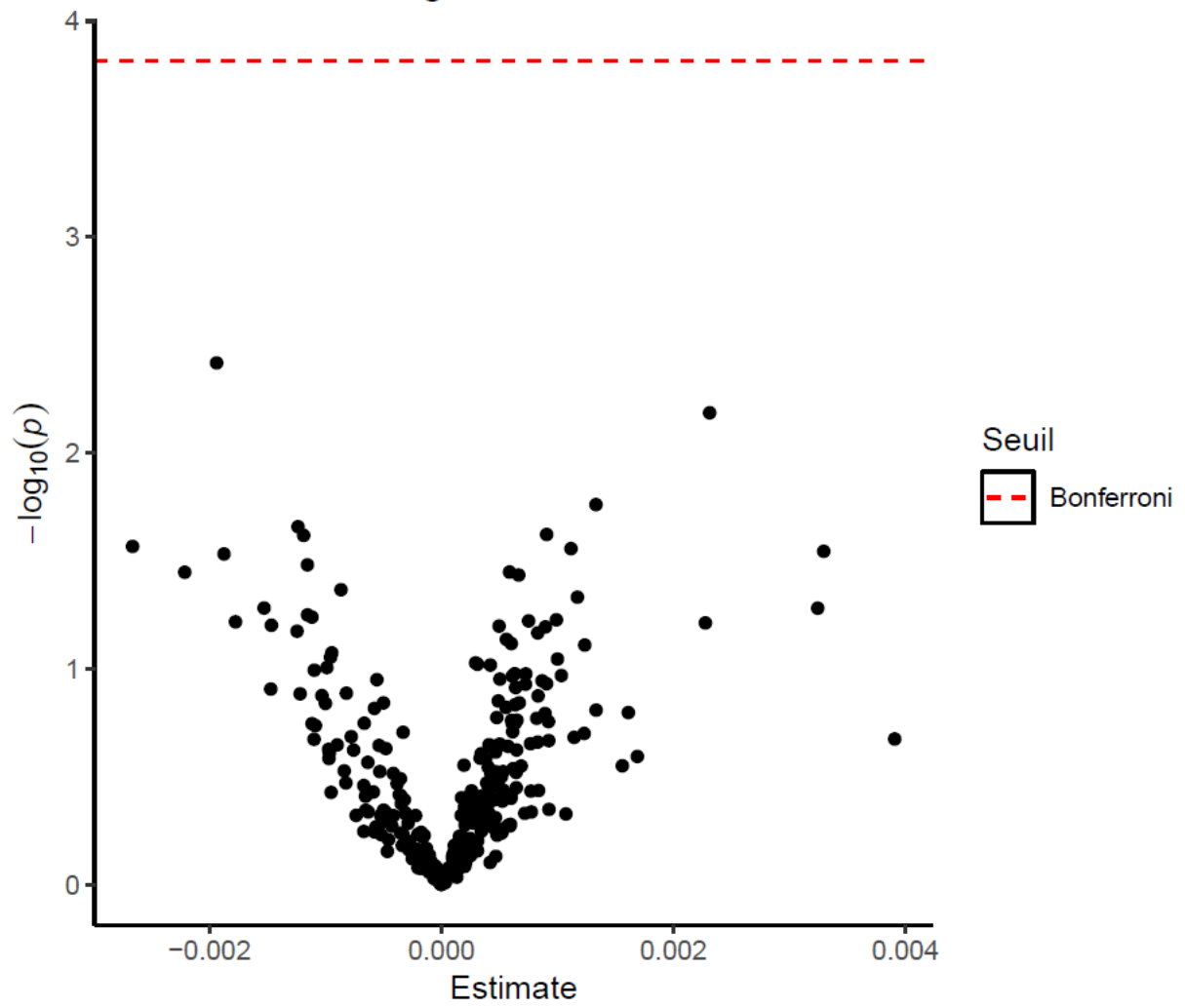


Figure III.D.3.46 : Volcano plot représentant les associations entre les gaz et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

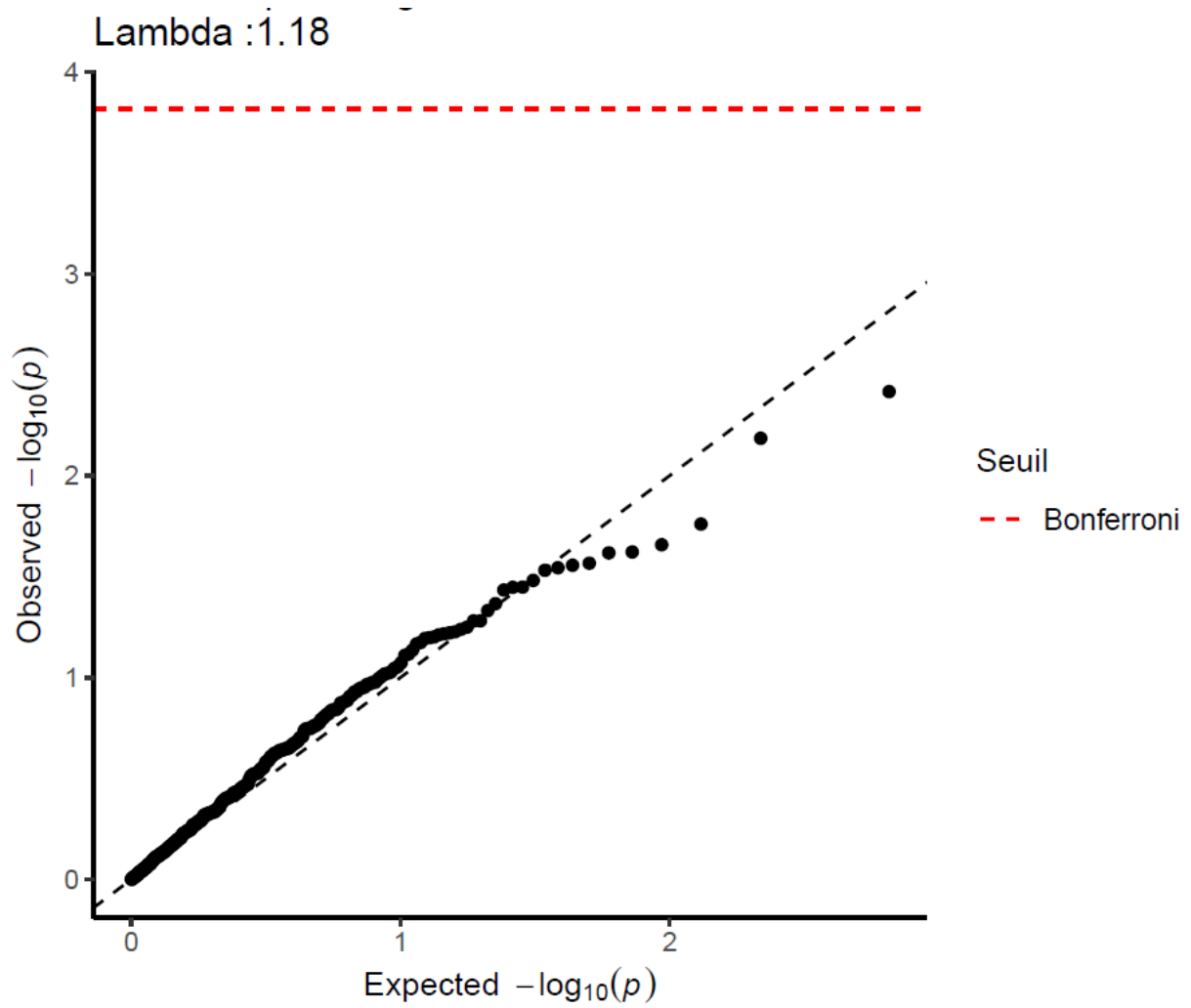


Figure III.D.3.47 : QQ plot représentant les associations entre les gaz et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

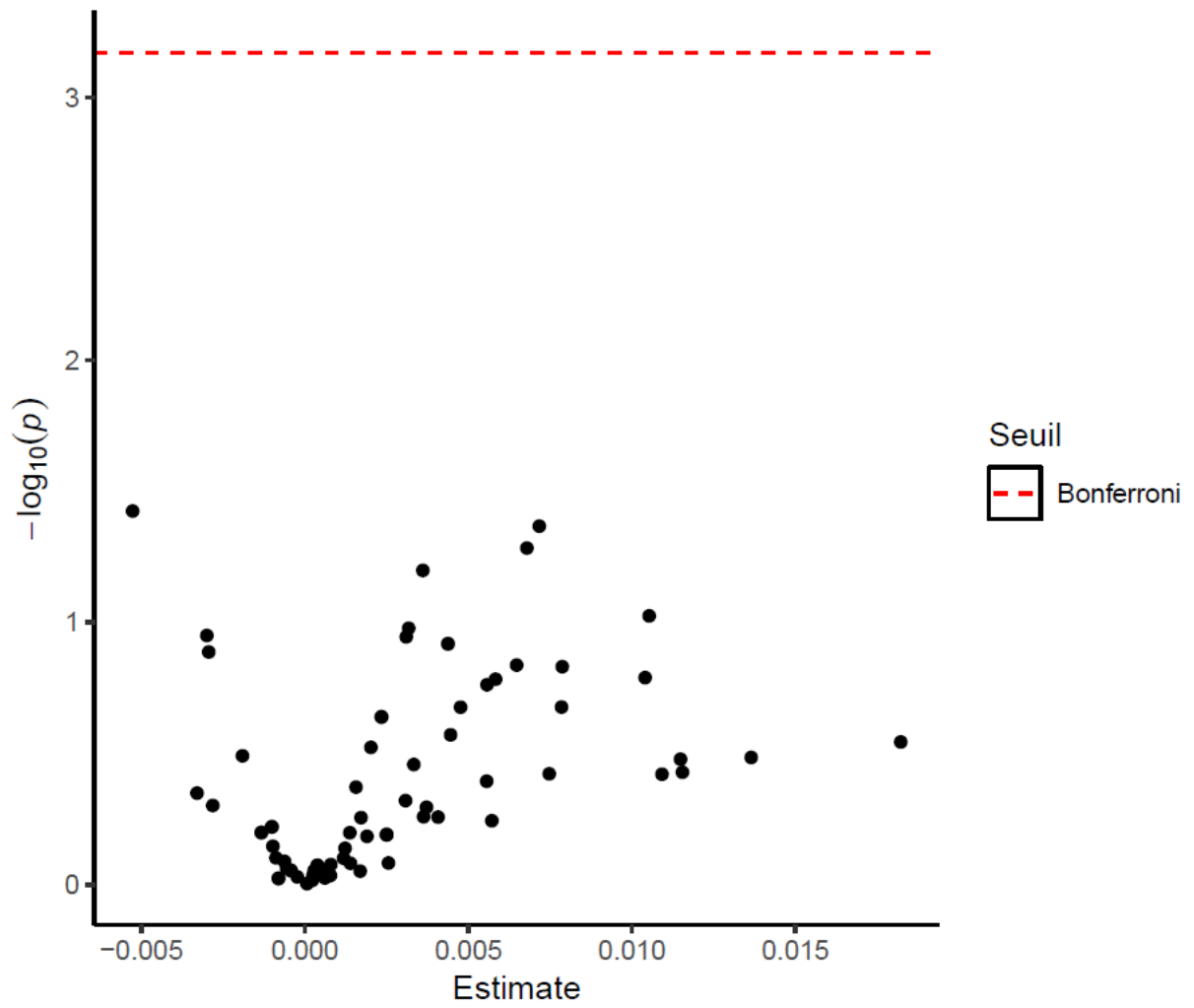


Figure III.D.3.48 : Volcano plot représentant les associations entre les gaz et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

Lambda :0.778
Likelihood Ratio – pvalue : 0.204

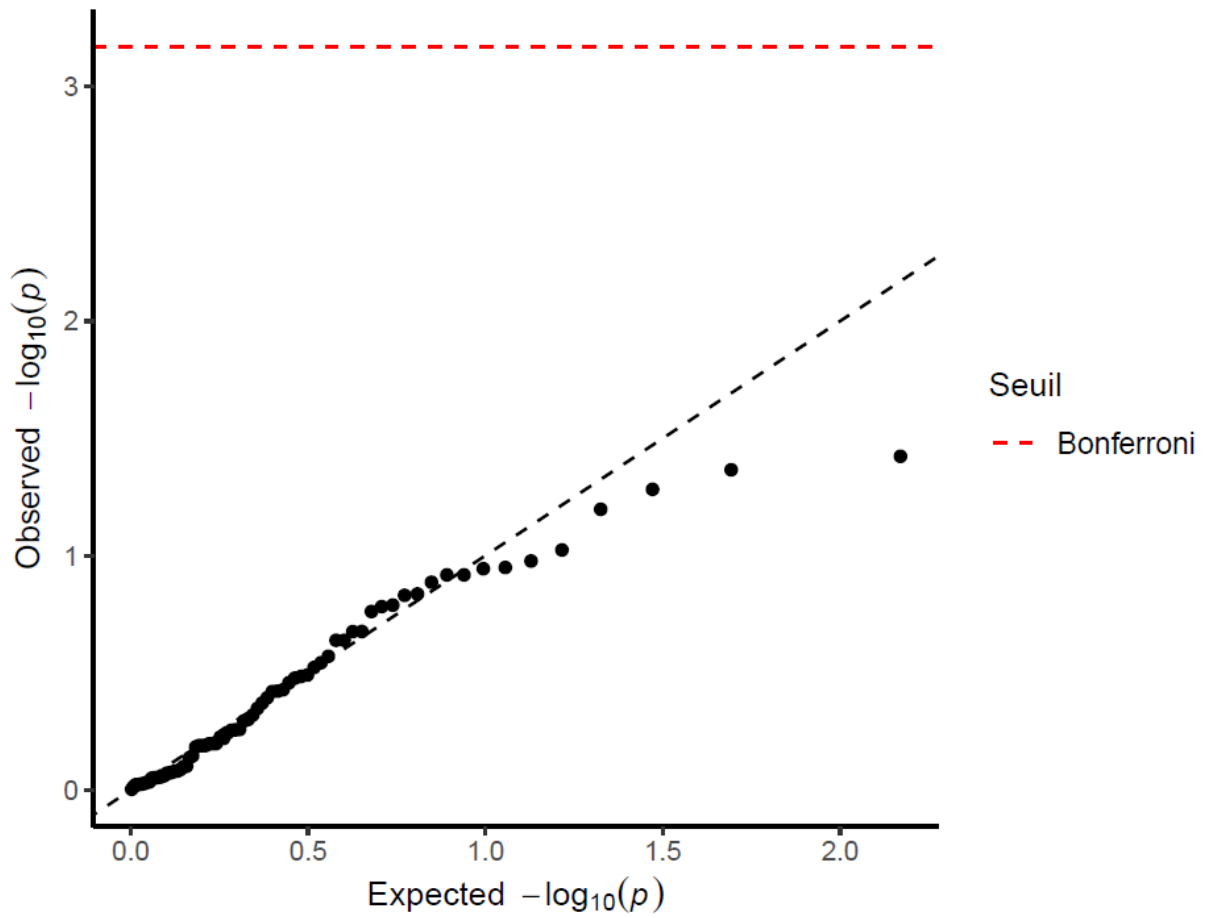


Figure III.D.3.49 : QQ plot représentant les associations entre les gaz et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

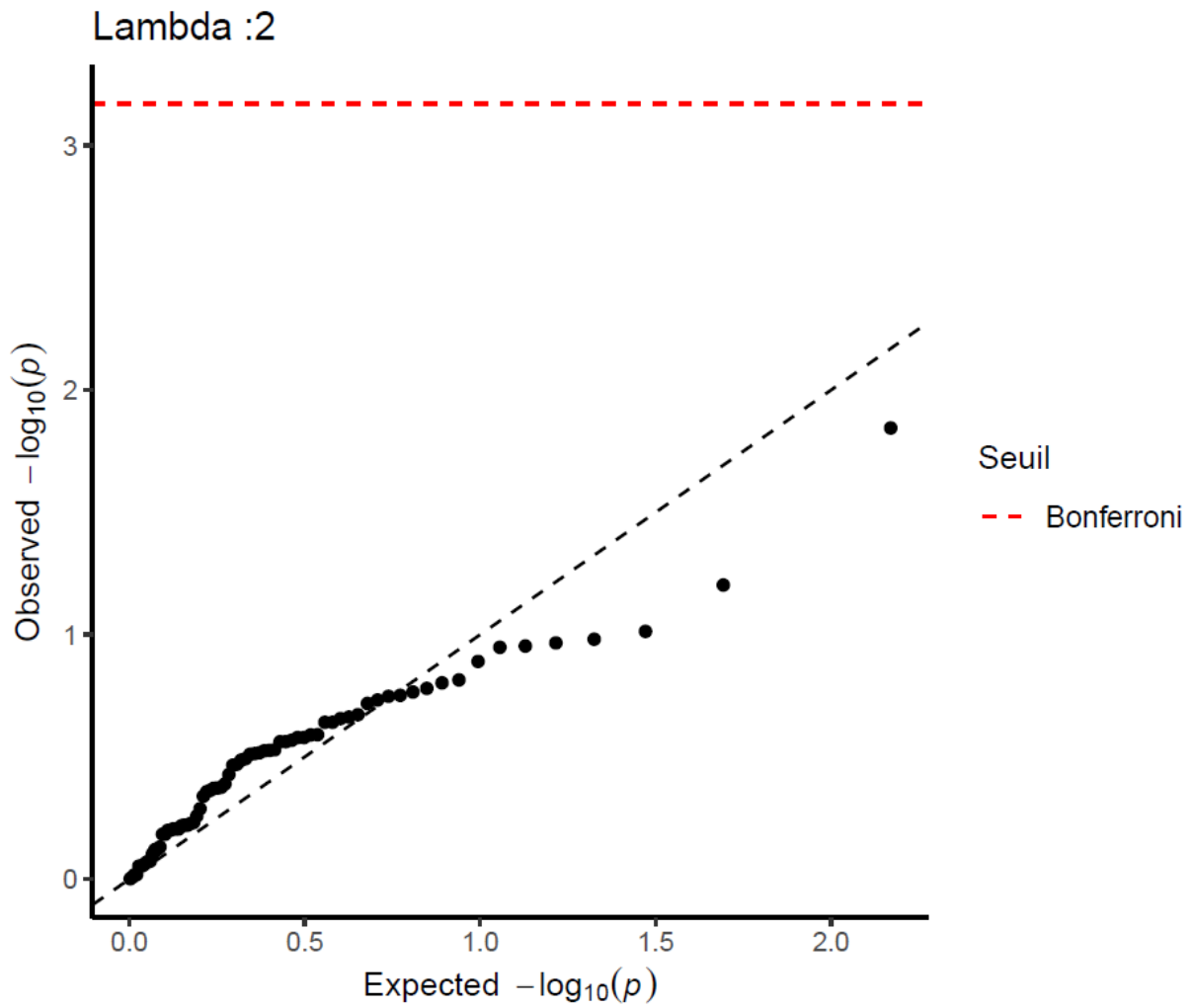


Figure III.D.3.50 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre les gaz et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe et la communauté urbaine

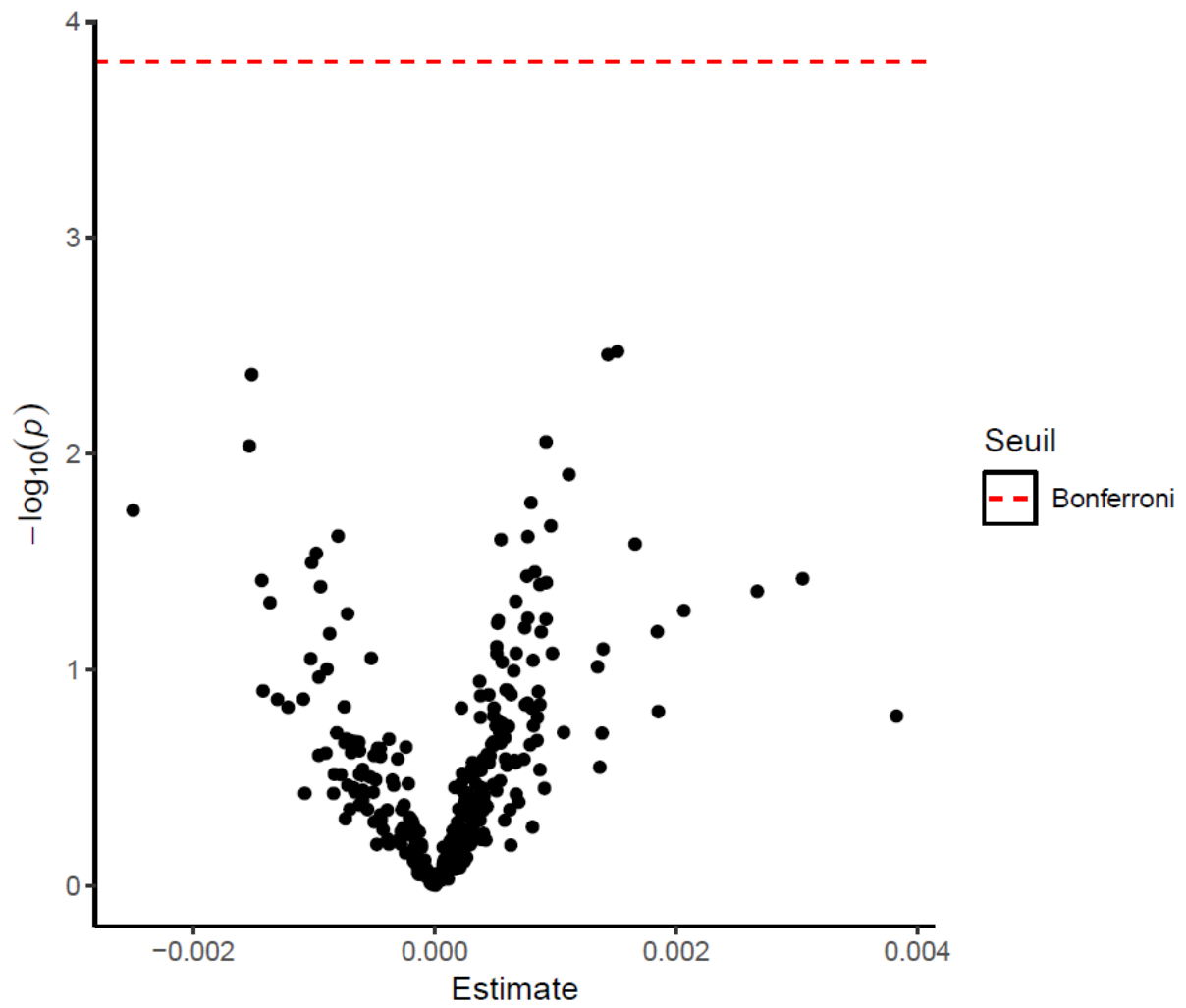


Figure III.D.3.51 : Volcano plot représentant les associations entre les fumées et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

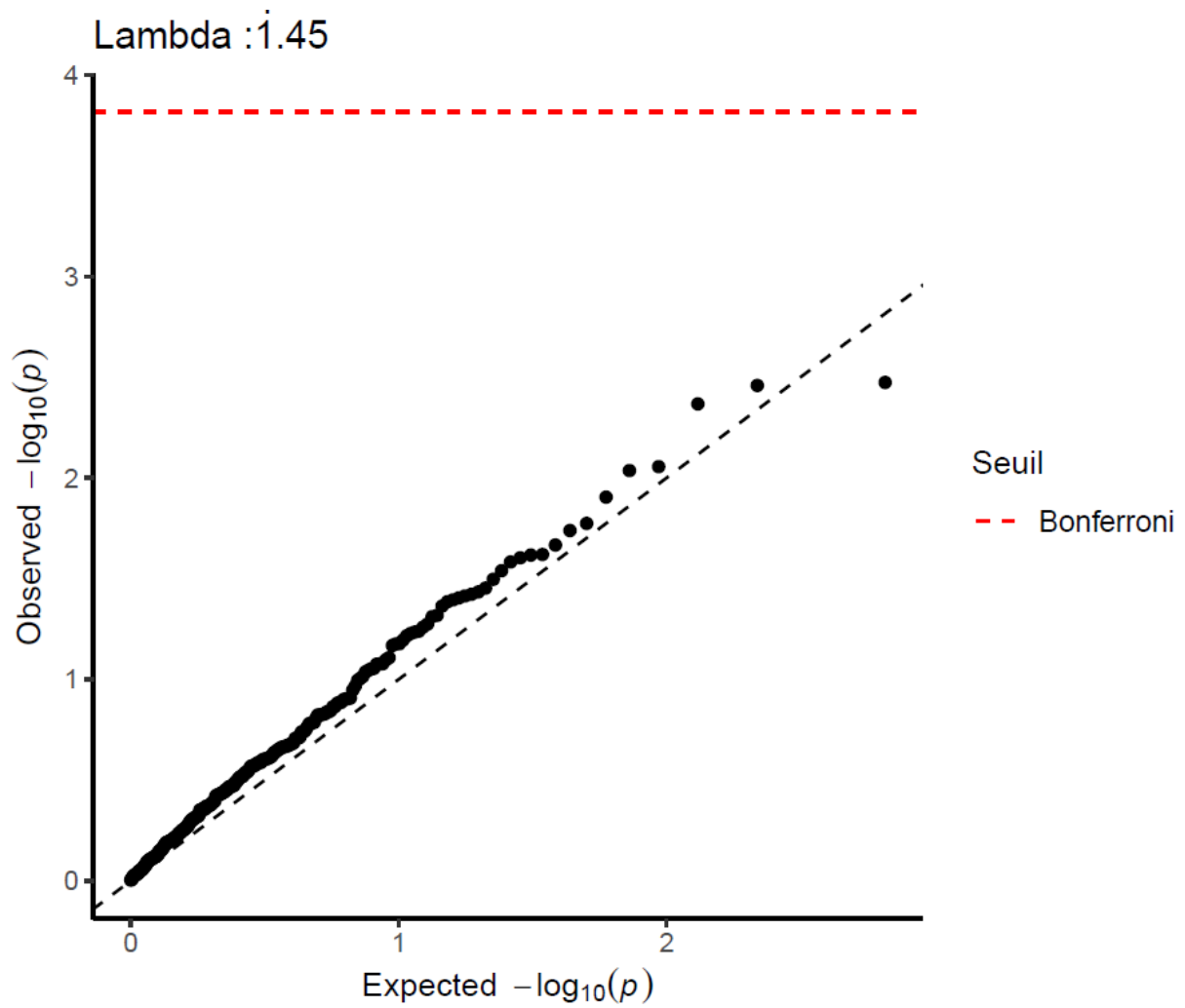


Figure III.D.3.52 : QQ plot représentant les associations entre les fumées et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine

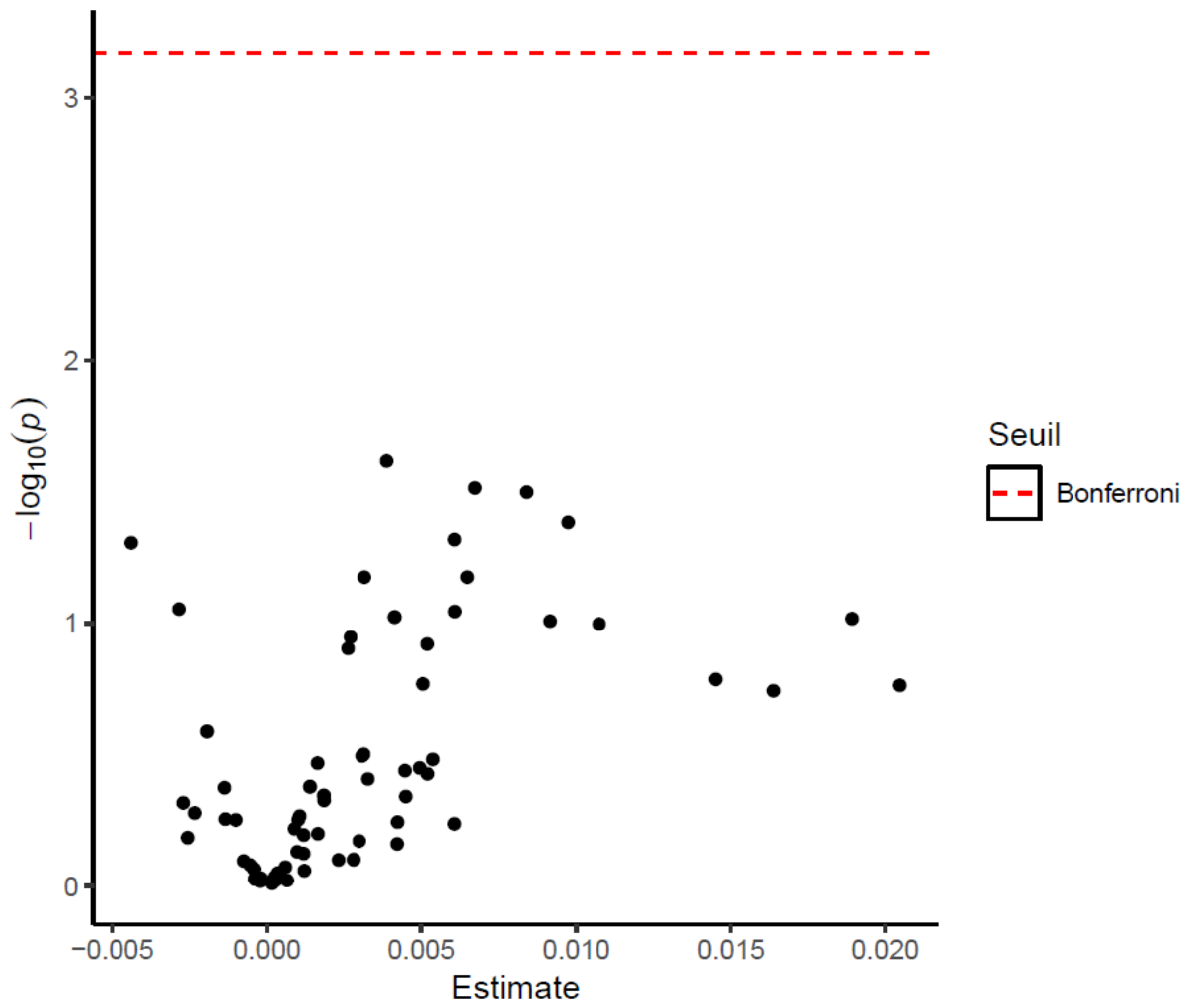


Figure III.D.3.53 : Volcano plot représentant les associations entre les fumées et **les voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

Lambda : 1.24
Likelihood Ratio - pvalue : 0.0342

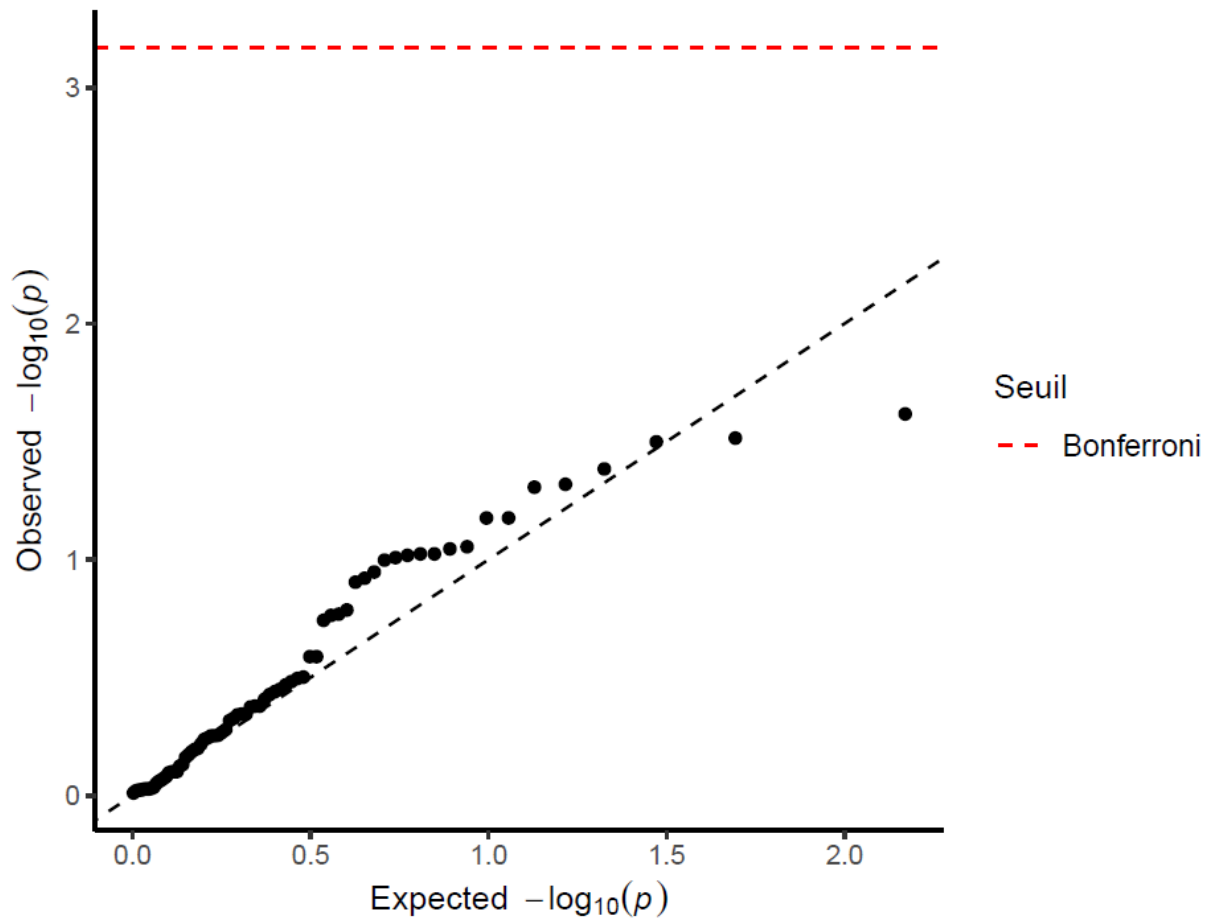


Figure III.D.3.54 : QQ plot représentant les associations entre les fumées et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

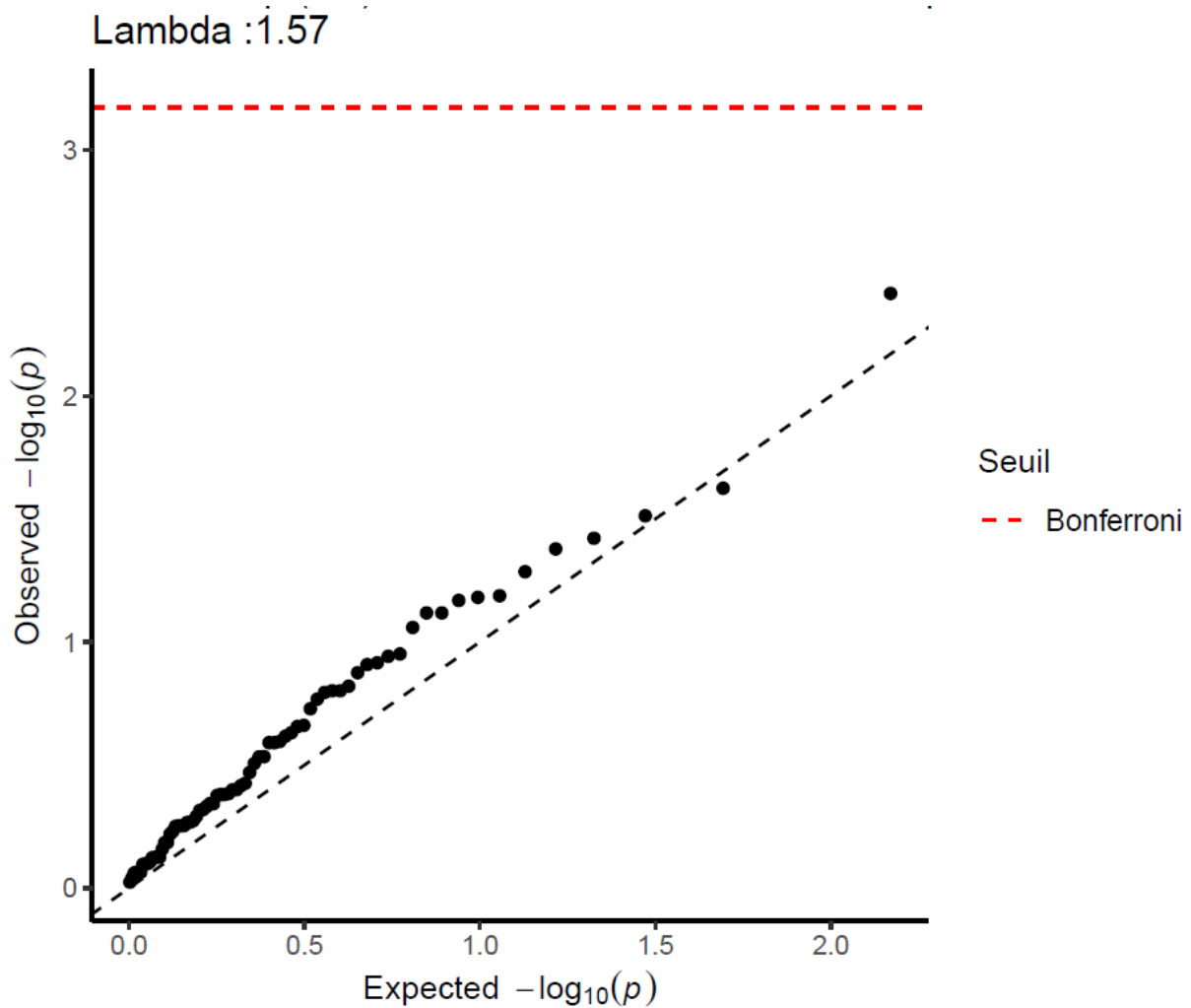


Figure III.D.3.55 : QQ plot représentant les p des rapports de vraisemblance des associations entre les fumées et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe et la communauté urbaine

(5) Métaux lourds : cadmium

L'exposition au cadmium n'était pas significativement associée aux métabolites (figures III.D.3.56 et III.D.3.57) ni aux voies métaboliques (figures III.D.3.58, III.D.3.59 et III.D.3.60). En revanche, le rapport de vraisemblance sur l'ensemble des voies métaboliques était significatif ($p = 0,0151$) (figure III.D.3.59), suggérant un impact du cadmium sur au moins une voie métabolique.

Nous avons également regardé les associations avec l'exposition au plomb, les résultats sont développés dans la partie correspondante (E. Associations entre plombémie et métabolites et voies métaboliques).

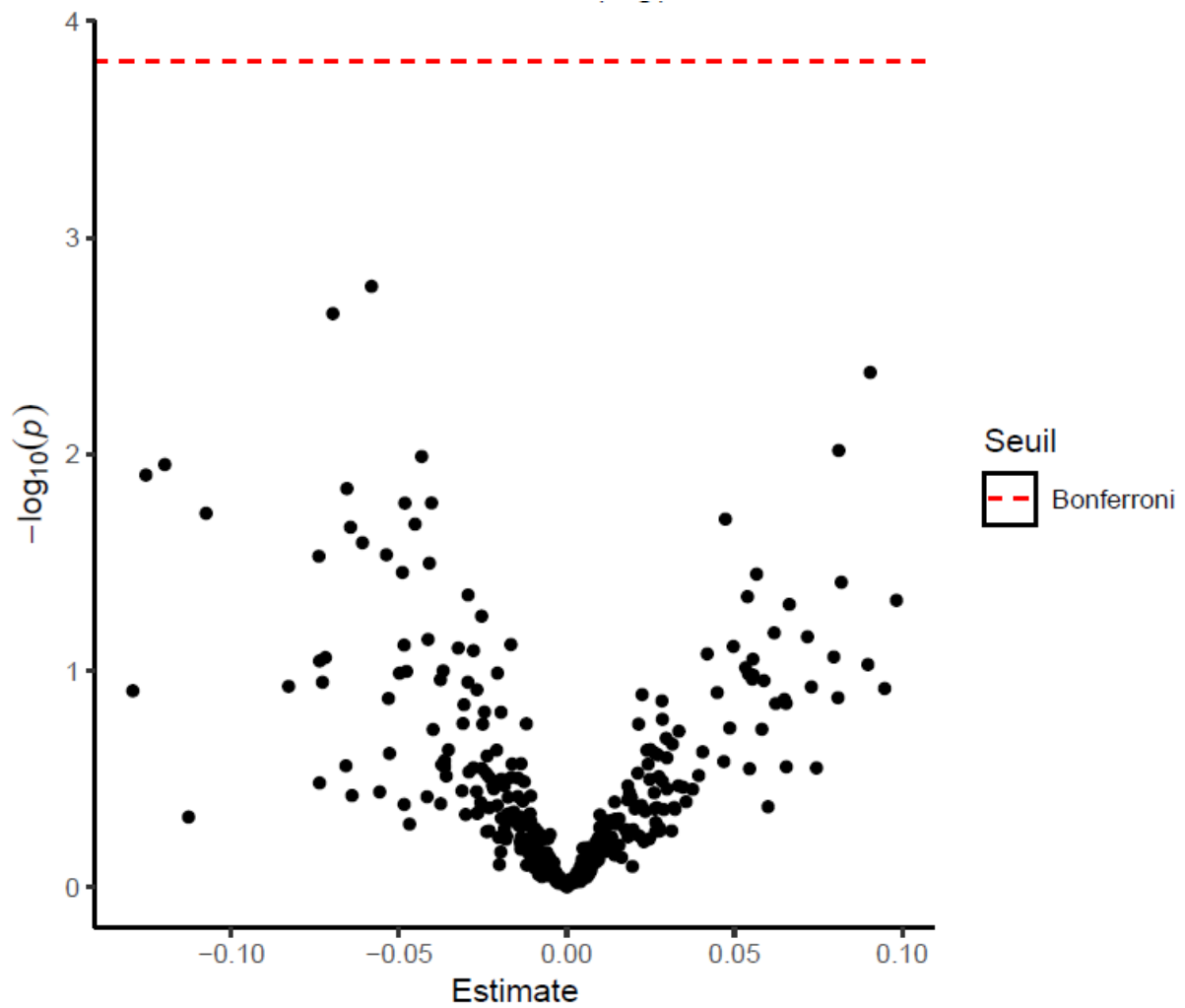


Figure III.D.3.56 : Volcano plot représentant les associations entre la concentration sanguine en cadmium et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

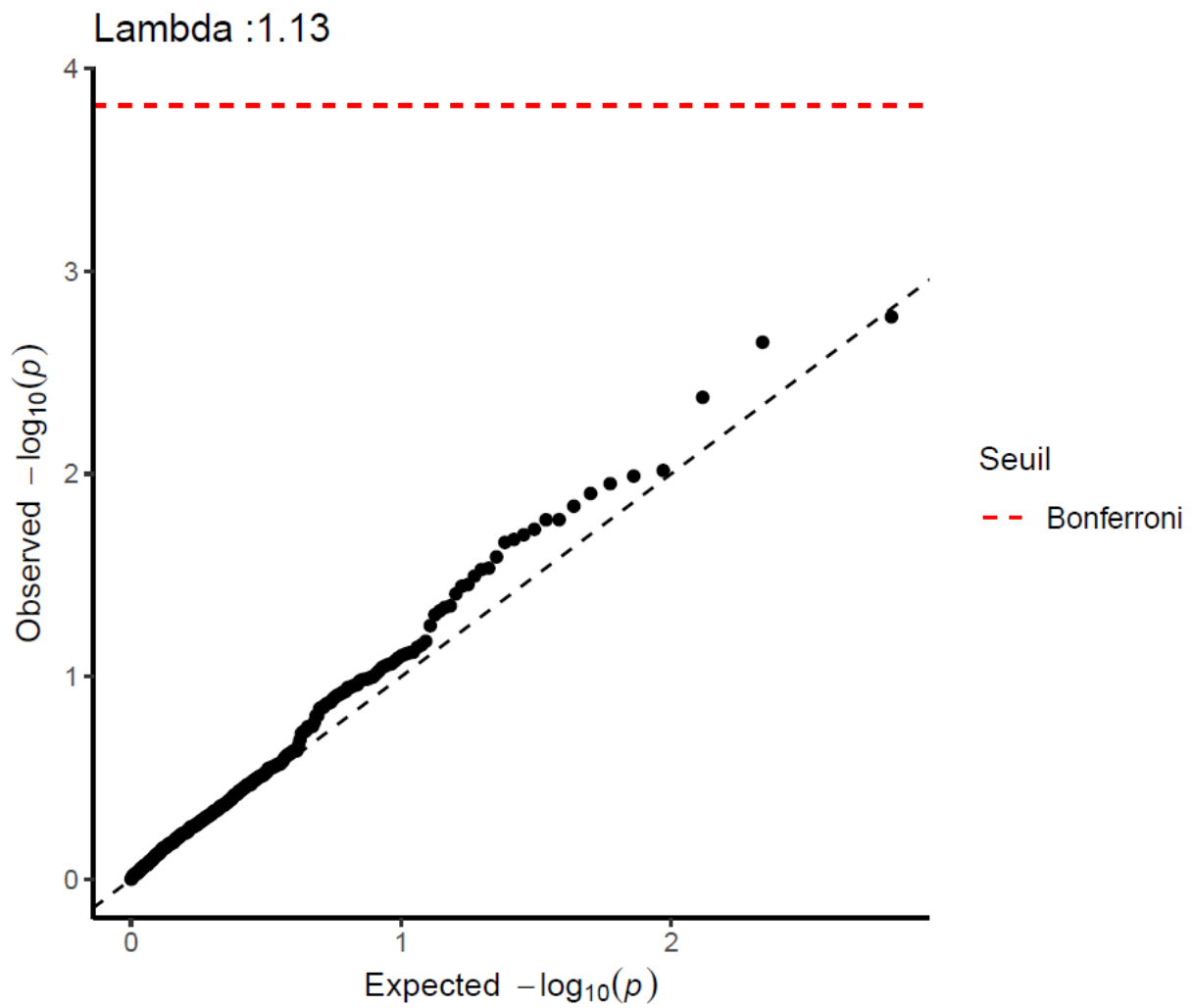


Figure III.D.3.57 : QQ plot représentant les associations entre la concentration sanguine en cadmium et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

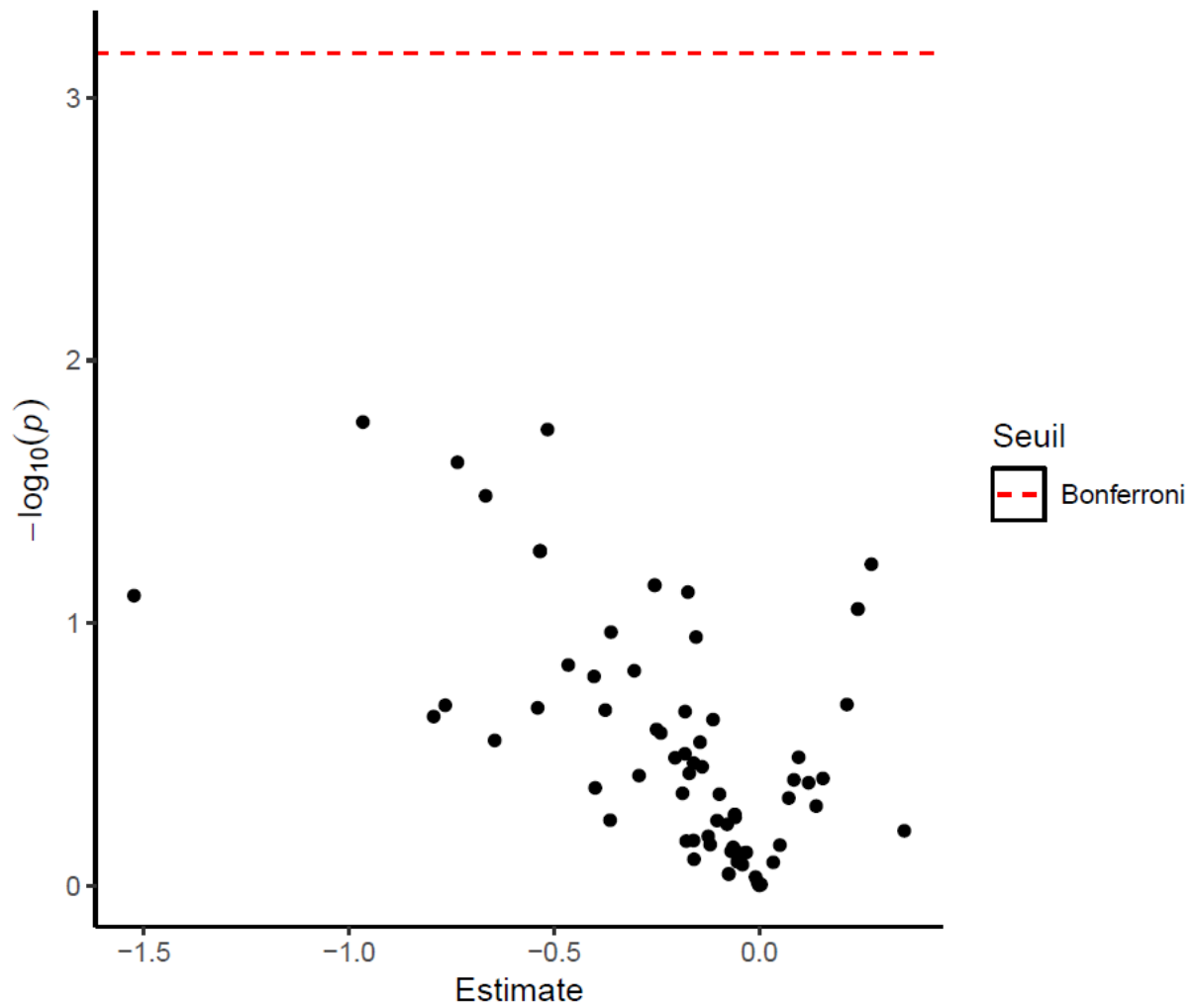


Figure III.D.3.58 : Volcano plot représentant les associations entre la concentration sanguine en cadmium et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

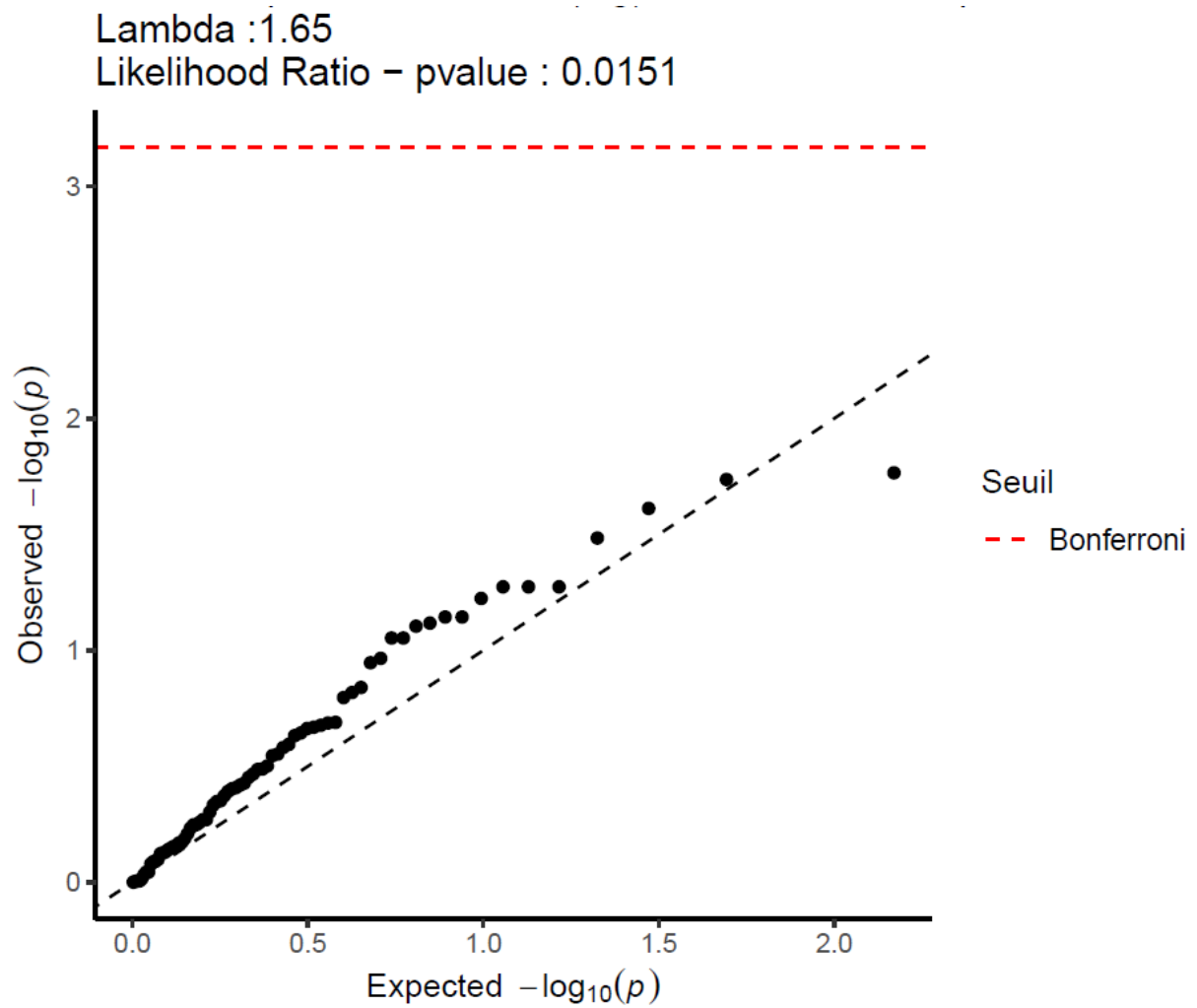


Figure III.D.3.59 : QQ plot représentant les associations entre la concentration sanguine en cadmium et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

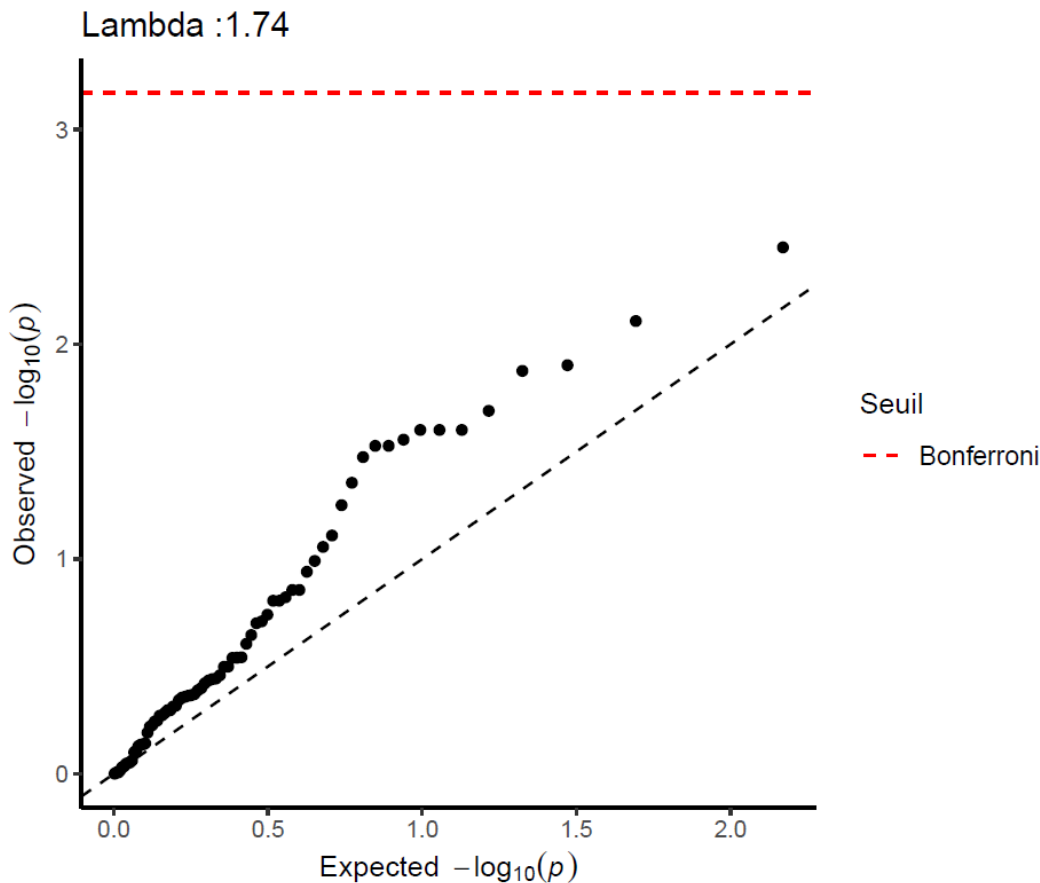


Figure III.D.3.60 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre la concentration sanguine en cadmium et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

4. Discussion de ces résultats

Les analyses sur des variables de démonstration ont permis de retrouver des associations attendues, suggérant la capacité de l'analyse statistique à identifier les effets métaboliques d'expositions ayant un impact important. Nous avons retrouvé des associations entre l'exposition aux PM₁₀ à court terme et un métabolite, entre l'exposition professionnelle aux vapeurs et 3 métabolites et entre l'exposition professionnelle aux poussières et 3 voies métaboliques. Pour les autres expositions environnementales, les associations étaient non significatives. Le tableau III.D.4.1 est une synthèse des résultats.

Nous avons donc retrouvé peu d'association avec les variables environnementales d'intérêts. La méthode a démontré son efficacité pour identifier les modifications métaboliques pour des expositions relativement importantes et/ou facilement mesurables. Les expositions environnementales observées dans notre population restent heureusement à des niveaux modérés. De plus, les mesures d'expositions sont souvent indirectes (modélisation de la pollution atmosphérique par exemple), ce qui peut conduire à une imprécision de la mesure (biais non différentiel diminuant la puissance de l'étude). Ces éléments, ainsi que le possible relativement faible impact de ces expositions sur le métabolisme, pourraient expliquer le peu d'associations identifiées avec certaines variables environnementales. Ces éléments sont détaillés dans la partie Discussion générale.

Tableau III.D.4.1 : Synthèse

Exposition	Métabolomique	Explication physiopathologique
Consommation de fruits	Nombreux métabolites associés et nombreuses voies métaboliques associées	Impact majeur de la consommation de fruits sur le métabolisme attendu
Sexe	Nombreux métabolites associés et nombreuses voies métaboliques associées	Différences importantes entre les hommes et les femmes au niveau du métabolisme attendues
Prise d'aspirine	3 métabolites associés	1 métabolite associé à l'aspirine dans les médicaments, 1 métabolite connu de l'aspirine et hypothèse difficile pour le troisième métabolite
PM ₁₀ court terme	1 métabolite associé, aucune voie métabolique associée	Hypothèse sur le stress oxydant, analyse plus approfondie à réaliser
PM ₁₀ long terme	NS	Non concerné
NO ₂ court terme	NS	Non concerné
NO ₂ long terme	NS	Non concerné
O ₃ court terme	NS	Non concerné
Moisissures	NS	Non concerné
Vapeurs	3 métabolites associés	Hypothèses difficile à formuler actuellement
Poussières	3 voies métaboliques associées	Hypothèse sur le stress oxydant, analyse plus approfondie à réaliser
Gaz	NS	Non concerné
Fumées	NS	Non concerné
Cadmium	NS	Non concerné
Plomb	22 métabolites associés, 50 voies métaboliques associées	Voies liées principalement au glutathion, impliqué dans le stress oxydant

NS : résultat non significatif

E. Associations entre plombémie et métabolites et voies métaboliques

1. Introduction

a) Exposition au plomb

Le plomb est un élément lourd toxique très répandu dans l'environnement (43,264). Bien que l'année 2000 ait marqué la fin du plomb-tétraéthyle dans l'essence en France (265) (et en 2021 la fin du plomb-tétraéthyle dans l'essence, partout dans le Monde (266)), l'intoxication au plomb reste un enjeu majeur de santé publique.

Les principales sources de contamination sont l'ingestion d'aliments ou d'eau contaminés (par exemple la consommation de fruits et légumes de sols contaminés ou la consommation d'eau potable des canalisations en plomb), l'exposition professionnelle (par exemple, métallurgie ou recyclage de batteries) ou encore l'ingestion accidentelle de sol, de poussières ou de peintures au plomb. Par ailleurs, le plomb n'est pas biodégradable et s'accumule donc dans l'environnement (264). On observe donc les conséquences de pollutions anciennes.

Le plomb est rapidement absorbé dans la circulation sanguine. La demi-vie du plomb dans le sang est d'environ 30 jours (43,267). La concentration sanguine en plomb reflète la quantité de plomb circulant dans l'organisme mais ne permet pas d'évaluer la quantité stockée dans l'organisme. Lors d'une exposition prolongée sur plusieurs années, l'élimination du plomb est beaucoup plus lente en raison de l'accumulation de plomb dans les os, relargué progressivement dans la circulation sanguine (264). Ainsi, chez les adultes, la plombémie peut refléter une exposition récente mais peut aussi témoigner d'une exposition ancienne (relargage lent).

La méthode de référence pour évaluer l'exposition au plomb dans l'organisme est la mesure de la concentration sanguine en plomb (264,268).

(1) Les organes touchés par la toxicité du plomb :

Chez l'adulte, pour des expositions modérées à élevées, les effets toxiques du plomb sont connus et concernent presque tous les organes, entraînant des conséquences cliniques connues.

Système nerveux :

Chez l'adulte, des effets cliniques neurologiques ont été observés pour des doses relativement élevées (à partir de 250 µg/L dans le sang). Les manifestations neurologiques sont un temps de réaction allongé, des difficultés de concentration, ou encore des maux de tête (264). La neuropathie périphérique est également un symptôme fréquent chez les adultes (269).

Des manifestations neurologiques très graves apparaissent à des concentrations en plomb très élevées (1 000 µg/L chez l'adulte), et comprennent l'encéphalite, une augmentation de la pression intracrânienne, un coma ou des convulsions (264).

Système cardiovasculaire :

L'exposition au plomb est responsable d'effets délétères au niveau cardiovasculaire entraînant, par exemple, une hypertension artérielle, un risque d'AVC ou de cardiopathie coronaire (43,264).

Système rénal :

L'exposition au plomb peut entraîner également une néphrotoxicité pouvant conduire à des néphropathies aiguës ou chroniques, ainsi qu'à des cancers du rein (43,264).

Système squelettique : L'os est le principal site de stockage du plomb, la demi-vie du plomb dans l'os est de plusieurs décennies (267). Ainsi, le plomb s'accumule durablement dans les os, avant d'être relargué progressivement dans la circulation sanguine. Par ailleurs le plomb affecte les ostéoblastes, les ostéoclastes et les chondrocytes (43). Il entraîne une hypocalcémie, réduit l'épaisseur corticale de l'os et la densité osseuse, augmentant le risque de fracture osseuse (270).

Système reproducteur :

La toxicité du plomb s'observe également au niveau de la fertilité masculine et féminine. Chez l'homme, l'exposition au plomb est responsable d'une diminution du nombre et du volume des spermatozoïdes, d'une altération de leur motilité et de leur morphologie. Chez la femme, on observe des fausses couches, des naissances prématurées, un faible poids de naissance des nouveau-nés, des troubles du développement dans l'enfance (le plomb du sang maternel passe dans le placenta et le lait maternel) (43,264).

Système hématopoïétique :

Le plomb est responsable d'une toxicité au niveau de la moelle osseuse (271). Par ailleurs, le plomb inhibe la synthèse de l'hémoglobine, en bloquant des enzymes clés impliquées dans la synthèse de l'hème et réduit la durée de vie des érythrocytes en provoquant une déstabilisation et une fragilisation de leur membrane cellulaire, entraînant directement une anémie (43,264).

Système digestif :

Le plomb peut provoquer des atteintes digestives comme des nausées ou des douleurs abdominales (264).

(2) Les valeurs de références :

En France, plusieurs indicateurs différents existent.

Tout d'abord, les indicateurs biologiques d'exposition (IBE) qui évaluent les expositions des travailleurs dans un but de prévenir le risque sanitaire. Ils permettent d'établir des valeurs limites biologiques (VLB) en-dessous desquelles le risque est considéré comme raisonnable pour les travailleurs. C'est donc un seuil de risque par rapport aux effets sanitaires estimés.

Les VLB de plombémie pour les expositions professionnelles sont de 400 µg/L chez l'homme et de 300 µg/L chez la femme (272). En 2019, l'Anses a proposé la valeur de 180 µg/L comme valeur limite biologique (VLB) pour des effets neurocomportementaux (273).

Par ailleurs, il existe des valeurs biologiques de référence (VBR). En 2019, l'Anses a retenu des VBR à partir de l'étude nationale nutrition santé (ENNS) réalisée en 2006 dont les limites d'exposition correspondent à une plombémie de 60 µg/L pour les femmes et 85 µg/L pour les

hommes (273). Les valeurs biologiques de référence proviennent des niveaux d'imprégnation au plomb mesurés en population générale, elles correspondent au 95^{ème} percentile d'une étude en population générale (ESTEBAN) (39). Elles ont pour but d'aider à l'interprétation des niveaux d'exposition et non pas d'assurer un niveau de protection sanitaire.

Par ailleurs, d'après Santé Publique France, chez l'enfant, un niveau d'intervention rapide est fixé à une plombémie de 50 µg/L et un niveau de vigilance à 25 µg/L (39).

L'OMS recommande le seuil de plombémie à 50 µg/L comme seuil déclencheur d'un examen approfondi des sources d'exposition d'une personne au plomb, ainsi que de la mise en place d'actions visant à réduire ou à supprimer cette exposition (37,38). Toutefois, il est important de souligner que le plomb n'a aucun rôle physiologique dans l'organisme humain et qu'il n'y a pas de seuil en dessous duquel, l'absence de risque est certaine (264). En effet, des effets néfastes sur le cerveau ont été observées à de faibles niveaux d'exposition c'est-à-dire 50 µg/L et moins (37).

Cependant, les effets cliniques de faibles doses d'exposition chez les adultes ne sont pas clairement identifiés. Des perturbations métaboliques subtiles pourraient avoir des effets sur la santé à long terme. Ces altérations métaboliques, bien que, silencieuses, pourraient contribuer à des maladies chroniques, soulignant la nécessité d'étudier de manière plus approfondie les expositions plus faibles au plomb et leurs conséquences chez l'adulte.

De plus, à de faibles concentrations sanguines de plomb, les effets toxiques sont en grande partie subcliniques mais pourraient entraîner des conséquences à long terme sur le risque de pathologie chronique. Par conséquent, en l'absence de compréhension des impacts potentiels à long terme de l'exposition, les gouvernements peuvent être moins incités à agir (37).

b) Littérature

Dans la littérature, peu d'études ont été retrouvées concernant l'association entre l'exposition au plomb et la métabolomique. Certaines études ont été trouvées mais celles-ci concernaient des petits effectifs (< 100 individus) ou des populations particulières. Par exemple, une étude menée chez 99 femmes enceintes issues de la cohorte « Programming Research in Obesity, Growth, Environment, and Social Stressors birth » à Mexico a mesuré 215 métabolites sériques et a retrouvé des associations entre 31 métabolites et le plomb sérique (médiane (IIQ) à 2,9 (2,6) µg/dL). L'analyse d'enrichissement de voies métaboliques a montré des associations avec les classes de métabolites suivantes : acides gras et conjugués, acides aminés et peptides, purines (274). Une autre étude de métabolomique ciblée menée auprès de 95 individus au Vietnam (44 vivants dans un village près d'un site de recyclage de batteries au plomb contre 51 contrôles vivant dans un village de référence) a retrouvé 10 métabolites urinaires potentiellement biomarqueurs de l'exposition au plomb (mesurée par la plombémie). Les 44 individus étaient exposés à des doses élevées (moyenne à 48 µg/dL dans le groupe exposé contre 3,9 µg/dL dans le groupe de référence). Ces métabolites correspondaient au métabolisme de l'arginine et de la proline, au cycle de l'urée, à la biosynthèse de la citrulline, au métabolisme de l'histidine (glutamate, acide imidazole acétique), au métabolisme de la glycine et de la sérine, ainsi qu'au métabolisme de la porphyrine et de la chlorophylle (275).

D'autres études, interventionnelles, ont été retrouvées chez les animaux. Par exemple, une étude a étudié l'effet du plomb sur le microbiome intestinal de souris exposées au plomb (par rapport à un groupe témoins) (276). Une autre, a étudié les effets du plomb sur le métabolisme de la crevette (277).

En revanche, peu d'étude ont été retrouvées en population générale sur de larges échantillons. Nous avons retrouvé une étude menée auprès de 399 hommes issus de l'étude « *Veterans Affairs Normative Aging Study* » à Boston aux Etats-Unis, qui s'intéressait aux signatures métaboliques de l'exposition au plomb mesurée dans le sang et les ongles de pieds. Les niveaux d'exposition étaient les suivants : 17 individus avaient une plombémie à 0 µg/dL, 290 avaient une plombémie inférieure à 5 µg/dL, 81 entre 5 et 10 µg/dL et 11 entre 10 et 30 µg/dL). Parmi les 858 métabolites plasmatiques mesurés, 154 étaient significativement associés au plomb sanguin dont 11 restaient significatifs après correction des tests multiples. L'analyse de voies métaboliques a montré des associations avec le métabolisme des glycérophospholipides et le métabolisme des sphingolipides. Pour le plomb mesuré dans les ongles, 70 associations ont été retrouvées dont 4 après correction des tests multiples (278).

c) **Objectif**

L'objectif de la présente étude était d'explorer les associations entre l'exposition au plomb et les modifications des processus métaboliques dans une population faiblement exposée, afin de mieux comprendre les mécanismes biologiques sous-jacents aux faibles doses d'exposition ainsi que leurs effets potentiellement subcliniques.

2. **Méthode**

a) **Population, mesure de l'exposition, mesures des métabolites**

Les méthodes concernant la population, la mesure de l'exposition au plomb ainsi que la mesure des métabolites ont été décrit dans la partie Méthode générale.

b) **Analyses statistiques**

Nous avons calculé la proportion de variance expliquée par l'effet batch (lot d'analyse) à l'aide d'une ANOVA, afin de quantifier ce biais de mesure et vérifier qu'il n'était pas trop important pour être corrigé par un ajustement.

Nous avons ensuite exploré les associations entre les concentrations sanguines de plomb et les métabolites pris individuellement à l'aide de modèles de régression linéaire. Les modèles ont été ajustés sur l'âge, le sexe, la communauté urbaine d'étude (la communauté urbaine de Lille ou Dunkerque) et le lot d'analyse. Afin de tenir compte des tests multiples, nous avons appliqué une correction par le taux de fausses découvertes (False Discovery Rate, FDR). Ces analyses ont été réalisées uniquement sur le sous-ensemble des métabolites annotés. Les résultats des analyses sur l'ensemble des métabolites sont présentés en **annexe 5**.

Dans un second temps, nous avons conduit des analyses basées sur les voies métaboliques. Nous nous sommes appuyés sur la revue systématique, et cette approche (par voies métaboliques) semble particulièrement pertinente pour répondre à notre objectif visant à mieux comprendre les mécanismes biologiques impliqués dans les effets de l'exposition.

Pour cela, nous avons eu recours à différentes méthodes :

Premièrement, nous avons utilisé l'ORA avec l'outil MetExplore (279), une plateforme bio-informatique en ligne permettant d'identifier les voies significativement associées à l'exposition à partir d'une liste de métabolites (métabolites significatifs). Cet outil offre également la possibilité de visualiser le réseau des voies significatives.

Deuxièmement, nous avons appliqué une méthode d'enrichissement appelée Metabolite Set Enrichment Analysis (MSEA).

Troisièmement, nous avons utilisé une méthode par scores, dans laquelle le score de chaque voie correspond à la somme des Z-scores (c'est-à-dire les valeurs standardisées sur l'ensemble des échantillons, avec une moyenne de 0 et un écart-type de 1) des métabolites impliqués dans cette voie.

Enfin, nous avons utilisé une méthode basée sur un test du rapport de vraisemblance, comparant deux modèles ayant l'exposition environnementale (le plomb) comme variable explicative : l'un incluant uniquement les variables d'ajustement, et l'autre incluant en plus les concentrations des métabolites appartenant à une voie donnée (comme proposée dans la revue systématique).

L'ensemble des analyses statistiques a été réalisé avec le logiciel R (version 2025.05.0). Le seuil de significativité a été fixé à 0,05.

3. Résultats de cette étude

a) Population

Nous avons sélectionné aléatoirement 1 000 individus issus de l'étude ELISABET. Parmi eux, 27 ont été exclus en raison d'un retrait de consentement, à la suite de l'envoi d'une lettre d'information dans le cadre de la conformité au RGPD demandée par le promoteur de l'étude. De plus, 47 échantillons n'étaient pas exploitables. Au total, 926 individus ont finalement été inclus dans cette étude, dont 404 (43,63 %) hommes et 522 (56,37 %) femmes. Les caractéristiques des participants sont présentées dans le tableau III.E.3.1.

Tableau III.E.3.1. Caractéristiques des participants de l'étude

	N = 926
Age (années) (moyenne (ET))	53,32 (7,28)
Sexe (hommes) (n (%))	404 (43,63)
Communauté urbaine (n (%))	
Lille	472 (50,97)
Dunkerque	454 (49,03)
Plomb ($\mu\text{g/L}$) (médiane (IIQ))	17,70 [12,20 - 25,40]

Abréviations : ET = écart type ; IIQ = intervalle interquartile

b) Exposition au plomb

Les analyses ont été réalisées dans 135 lots indépendants. La variance expliquée par l'effet de lot était de 25,13 %.

Les concentrations sanguines en plomb sont présentées dans le tableau III.E.3.1 et III.E.3.2. La concentration médiane était de 17,70 µg/L, avec un intervalle interquartile allant de 12,20 à 25,40 µg/L.

Le nombre de participants selon les catégories de concentration sanguine en plomb, pour les hommes et les femmes, est présenté dans le tableau III.E.3.2. Aucun des individus inclus dans l'étude ne dépassait les valeurs limites d'exposition professionnelle (400 µg/L chez les hommes et 300 µg/L chez les femmes). Toutefois, 6 hommes et 5 femmes présentaient des concentrations sanguines en plomb supérieures aux valeurs de référence biologiques pour les adultes (85 µg/L chez les hommes et 60 µg/L chez les femmes).

La grande majorité (901 participants, dont 386 hommes et 515 femmes) avait des valeurs inférieures au seuil d'alerte défini pour les enfants (50 µg/L).

Tableau III.E.3.2. Description de la concentration sanguine en plomb selon le sexe

	Hommes				
Concentration (µg/L)	(0,25]	(25,50]	(50,85]	(85,400]	> 400
Effectif	267	119	12	6	0
	Femmes				
Concentration (µg/L)	(0,25]	(25,50]	(50,60]	(60,300]	> 300
Effectif	416	99	2	5	0

25 µg/L : seuil d'alerte chez l'enfant

50 µg/L : seuil d'intervention immédiate chez l'enfant

60 µg/L : limite chez les femmes

85 µg/L : limite chez les hommes

300 et 400 µg/L : limite pour l'exposition professionnelle

c) Variables métaboliques

Comme décrit précédemment, 10 043 signaux ont été détectés et 327 métabolites ont pu être annotés avec un niveau de preuve suffisant.

Les métabolites annotés appartenait à 74 voies métaboliques différentes.

d) Analyses par métabolites individuels

Les analyses menées sur les métabolites pris individuellement ont révélé que 22 métabolites étaient significativement associés à l'exposition au plomb après correction par le FDR. Les résultats sont présentés dans le tableau III.E.3.3 et les figures III.E.3.1 et III.E.3.2.

Tableau III.E.3.3. Métabolites associés à la concentration sanguine au plomb ($\mu\text{g/L}$)

Métabolite	Coefficient d'association (en unité de surface de pics chromatographiques)	p (FDR)
Acide tartrique	0,72	< 0,001
Cystathionine	-0,27	0,0024
Nicotinamide N-oxide	-0,1	0,0028
Hydroxyméthyluracile	-0,17	0,0035
Méthyladénine	-0,082	0,014
Acide 3,5-dihydroxybenzoïque	-0,200	0,017
Cystine	-0,18	0,018
Acide 2-isopropylmalique	0,35	0,019
Acide abscissique	-0,21	0,020
Méthylpyridines	0,39	0,021
Acides N-acétyl-aminosalicyliques	0,32	0,021
Phénylacétaldéhyde	0,25	0,024
Acide isohomovanillique (acide hydroxy-méthoxyphénylacétique)/acide homovanillique	0,21	0,027
Trigonelline	0,188	0,028
N-acétyl-glycine	0,15	0,031
Acide malonique.	0,19	0,035
Acide méthylhippurique	-0,24	0,036
Acide indolylactique	0,22	0,037
Pentoses	0,16	0,037
Hydroxybenzaldéhyde	0,22	0,043
Acide quinique	0,24	0,048
Acide oxobutyrique	0,21	0,048

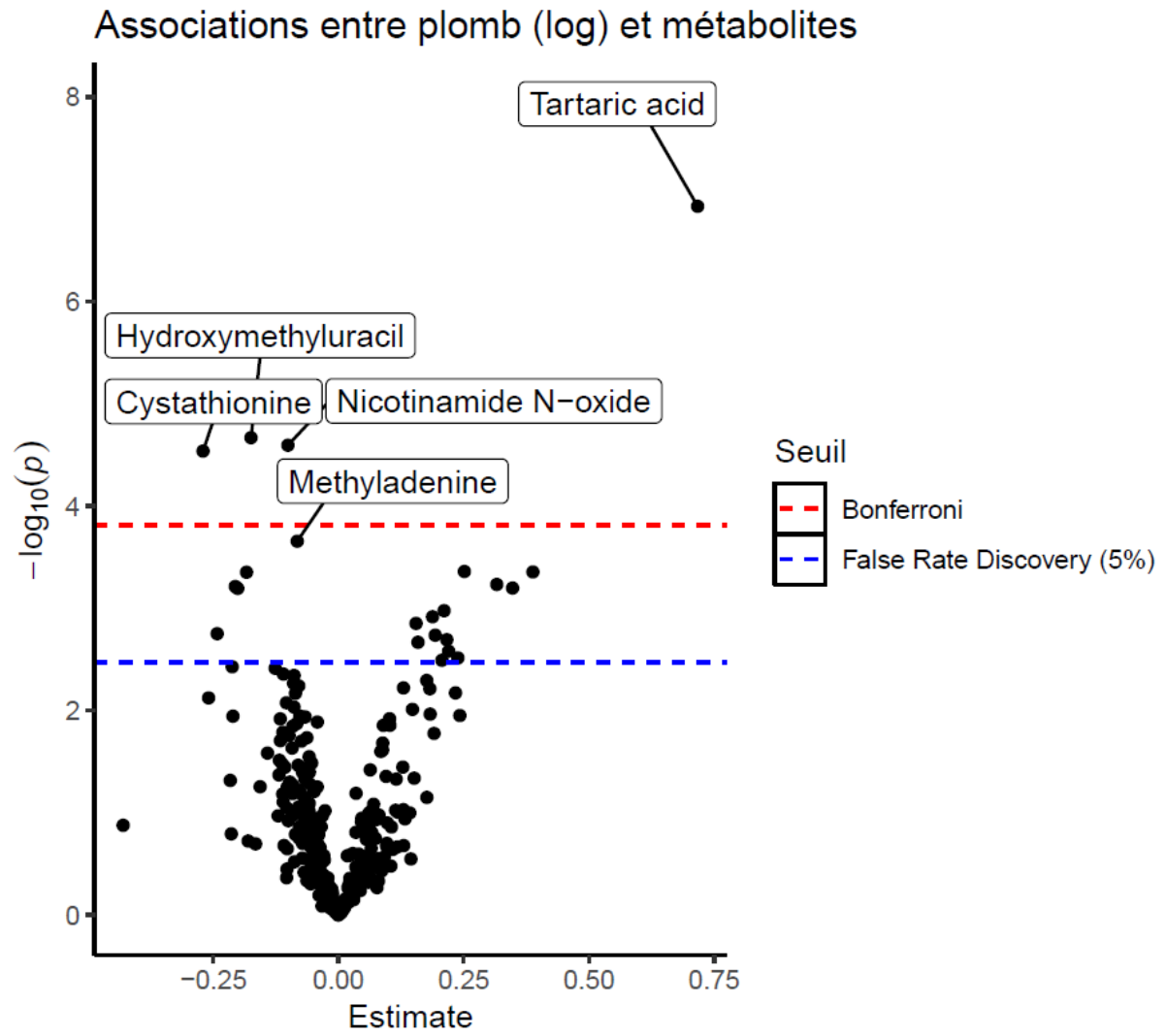


Figure III.E.3.1 : Volcano plot représentant les associations entre la concentration sanguine en plomb et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

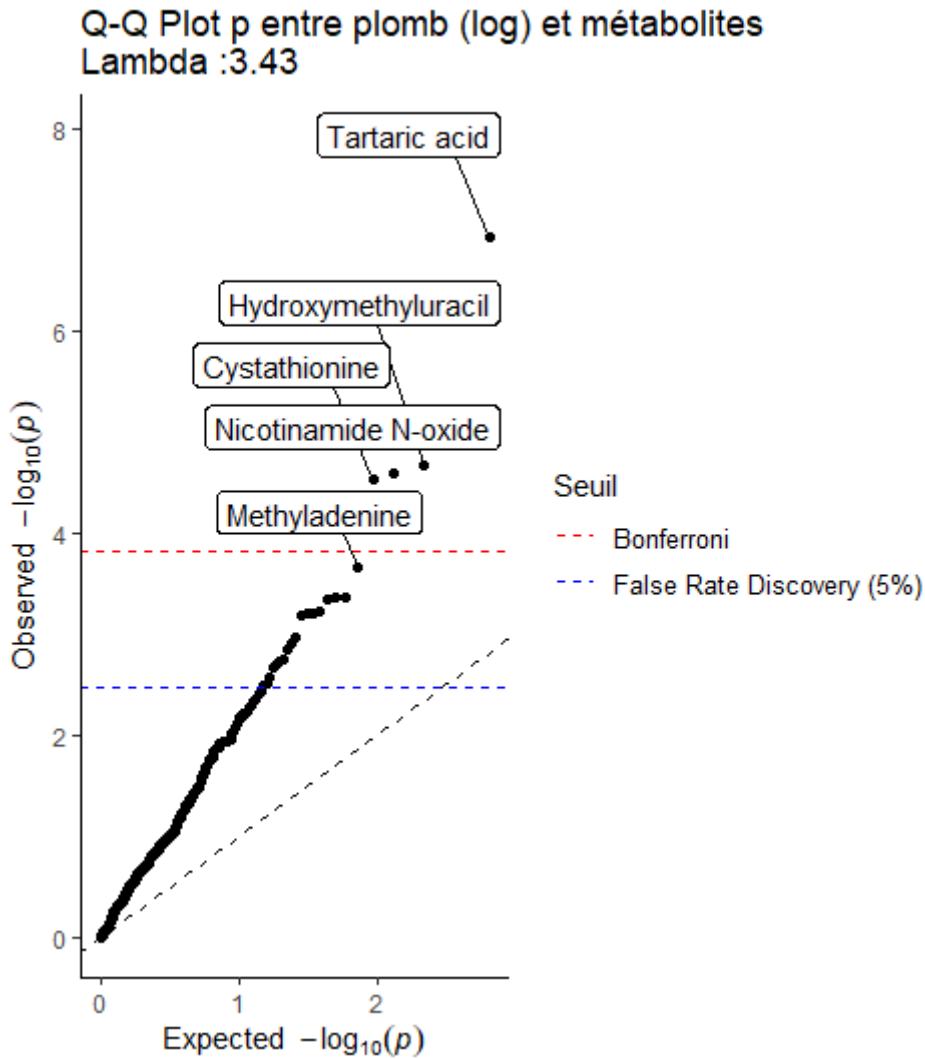


Figure III.E.3.2 : QQ plot représentant les associations entre la concentration sanguine en plomb et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

e) Analyses par voies métaboliques

(1) ORA (MetExplore)

Nous avons identifié 22 métabolites annotés significativement associés à l'exposition au plomb. Parmi eux, 16 disposaient d'un identifiant ChEBI disponible dans notre base de données. Après vérification dans MetaboAnalyst (280)— comme recommandé pour une utilisation conjointe avec MetExplore — et à l'aide de ChEBI (281), un métabolite s'est révélé avoir des identifiants ChEBI insuffisamment spécifiques (en raison d'isomères) et a donc été exclu. Ainsi, 15 métabolites ont été utilisés comme données d'entrée pour MetExplore.

Après correction par FDR, une voie métabolique était significativement associée aux concentrations sanguines de plomb, celle du métabolisme de la tyrosine.

(2) MSEA

En utilisant la méthode MSEA, nous n'avons trouvé aucune association significative entre l'exposition au plomb et les voies métaboliques.

(3) Z-score

En utilisant une méthode basée sur les Z-scores, nous avons identifié 18 voies métaboliques significativement associées aux concentrations sanguines de plomb après correction pour comparaisons multiples par FDR. Les résultats sont présentés dans le tableau III.E.3.4 ainsi que sur les figures III.E.3.3 et III.E.3.4.

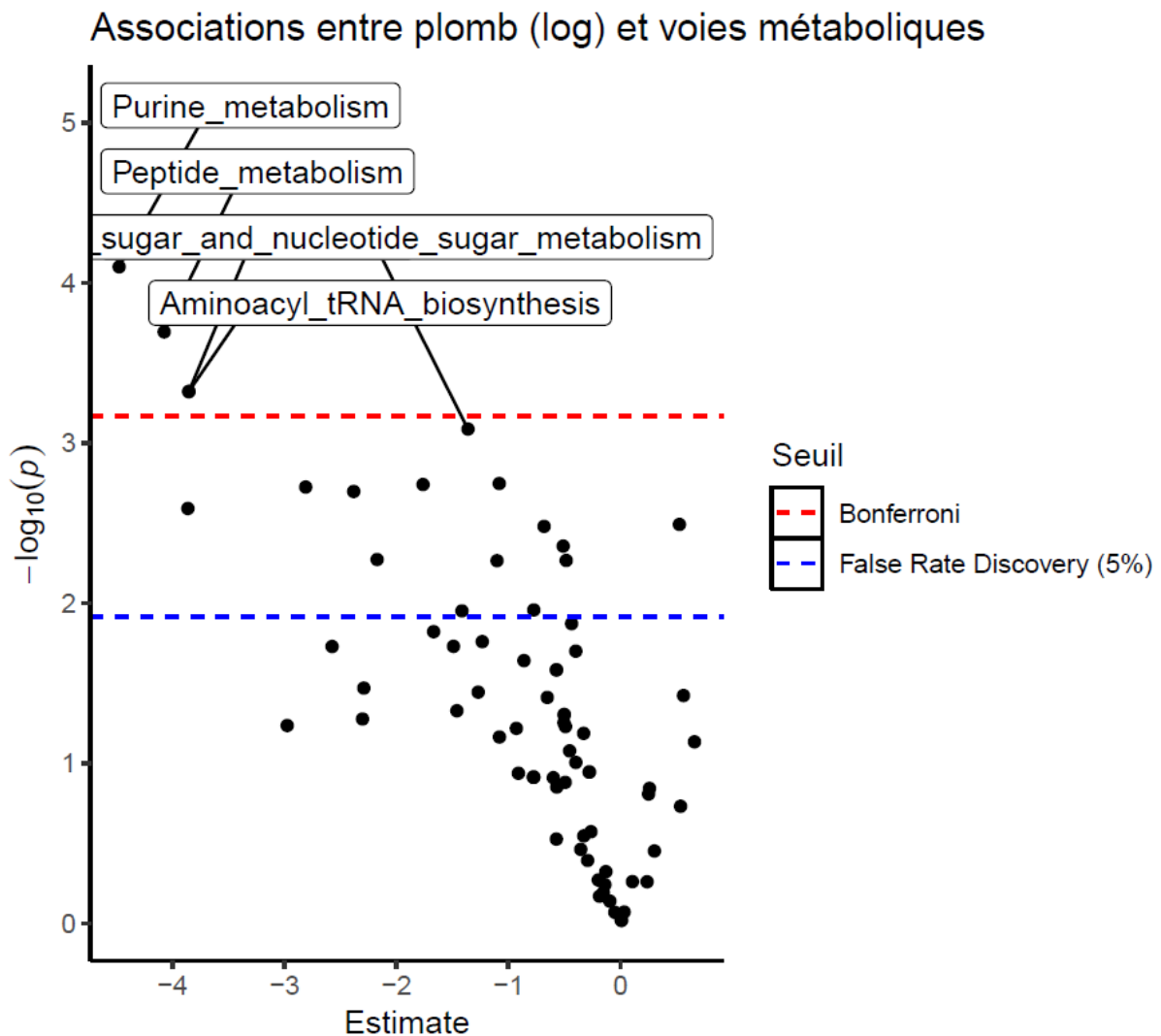


Figure III.E.3.3 : Volcano plot représentant les associations entre la concentration sanguine en plomb et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

Q-Q Plot p entre plomb (log) et voies métaboliques
 Lambda :7.87
 Likelihood Ratio – pvalue : 2.17e-07

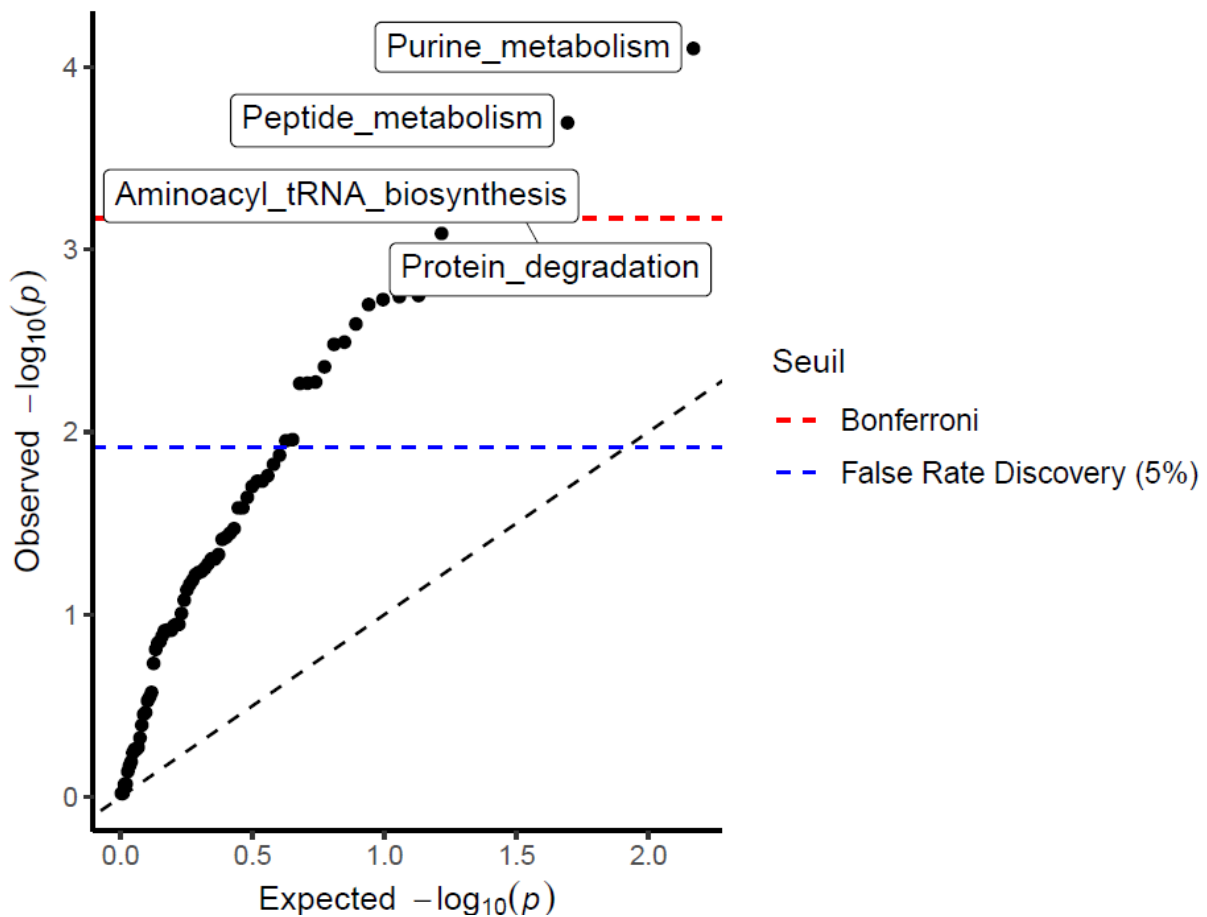


Figure III.E.3.4 : QQ plot représentant les associations entre la concentration sanguine en plomb et les **voies métaboliques (Z-score)**, ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

(4) Rapport de vraisemblance

La méthode du test du rapport de vraisemblance nous a permis d'identifier 50 voies métaboliques significativement associées à l'exposition au plomb (après ajustement par FDR). Les résultats sont présentés dans le tableau III.D.4 et sur la figure III.D.5.

Q-Q Plot p (RV) entre plomb (log) et voies métaboliques
 Lambda :13

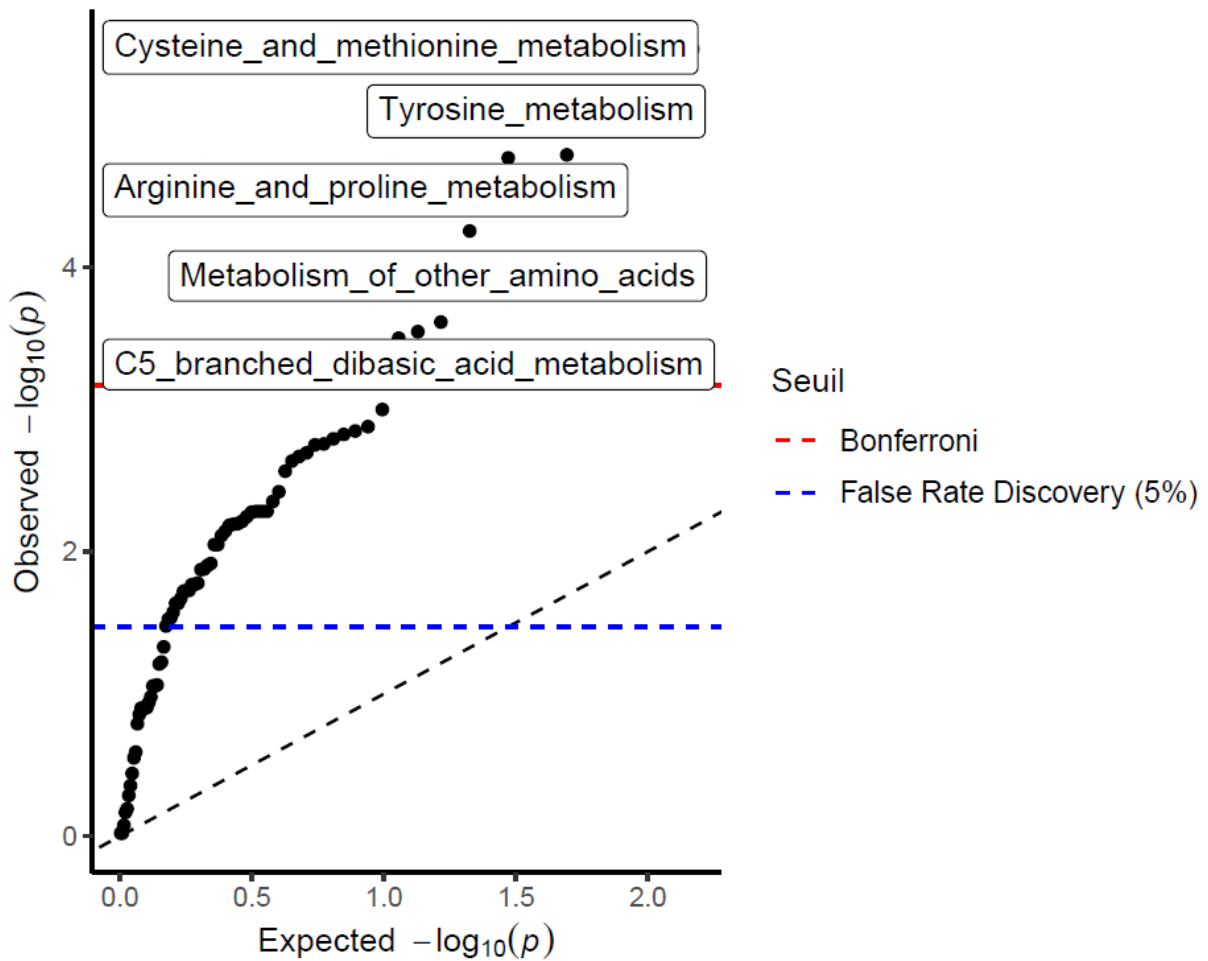


Figure III.E.3.5 : QQ plot représentant les p des **rapports de vraisemblance** des associations entre la concentration sanguine en plomb et les voies métaboliques (Z-score), ajustées sur l'âge, le sexe, la communauté urbaine et le lot d'analyse

(5) Comparaisons des voies métaboliques selon les méthodes

Les 18 voies métaboliques associées à la plombémie retrouvées avec la méthode par Z-score sont toutes sauf une (celle du métabolisme de la purine) retrouvées par la méthode du rapport de vraisemblance.

Le tableau III.E.3.4 présente les voies métaboliques identifiées par les différentes méthodes.

Tableau III.E.3.4. Comparaisons des voies métaboliques associées à la plombémie selon les méthodes statistiques

Voie métabolique	Rapport de vraisemblance		Méthode statistique		ORA	MSEA
	p (FDR)	Z-score	coefficient d'association (en écart-type de métabolite pour 1 µg/L de plombémie)	p (FDR)	p (FDR)	p (FDR)
Métabolisme de la cystéine et de la méthionine	< 0,001		-2,17	0,028	NS	NS
Métabolisme de l'arginine et de la proline	< 0,001			NS	NS	NS
Métabolisme de la tyrosine	< 0,001			NS	0,042	NS
Métabolisme des autres acides aminés	0,0010			NS	NS	NS
Métabolisme des glycérophospholipides	0,0033			NS	NS	NS
Métabolisme de la glycine, de la sérine et de la thréonine	0,0035			NS	NS	NS
Métabolisme des acides dibasiques ramifiés en C5	0,0036			NS	NS	NS
Biosynthèse des acides biliaires	0,0093			NS	NS	NS
Navette de la carnitine (réticulum endoplasmique)	0,0094			NS	NS	NS
Biosynthèse du squelette des terpénoïdes	0,0099		-0,51	0,025	NS	NS
Diverses voies	0,0099			NS	NS	NS
Métabolisme de l'histidine	0,010		-1,76	0,019	NS	NS
Métabolisme de l'alanine, de l'aspartate et du glutamate	0,010			NS	NS	NS
Métabolisme du glutathion	0,010		-1,082	0,022	NS	NS
Métabolisme du propanoate	0,010		0,53	0,022	NS	NS
Métabolisme du soufre	0,011		-1,10	0,025	NS	NS
Biosynthèse de la phénylalanine, de la tyrosine et du tryptophane	0,011		-3,86	0,019	NS	NS
Métabolisme des O-glycanes	0,011		-0,48	0,027	NS	NS
Voies isolées	0,015			NS	NS	NS
Métabolisme des sucres aminés et des nucléotides-oses	0,016		-1,36	0,012	NS	NS
Métabolisme des éicosanoïdes	0,016			NS	NS	NS
Métabolisme de la bêta-alanine	0,017			NS	NS	NS
Navette de la carnitine cytosolique	0,017			NS	NS	NS
Métabolisme de la thiamine	0,017		-0,68	0,020	NS	NS
Métabolisme de la lysine	0,017			NS	NS	NS
Métabolisme des pyrimidines	0,017		-2,81	0,017	NS	NS
Métabolisme du butanoate	0,017			NS	NS	NS
Navette de la carnitine peroxyosmale	0,018			NS	NS	NS
Métabolisme des sphingolipides	0,018			NS	NS	NS
Navette de la carnitine mitochondriale	0,018			NS	NS	NS
Métabolisme des acides gras oméga-3	0,018			NS	NS	NS
Synthèse de la coenzyme A	0,020			NS	NS	NS
Biosynthèse du pantothénate et de la coenzyme A	0,021			NS	NS	NS
Métabolisme du pyruvate	0,026			NS	NS	NS
Métabolisme des nucléotides	0,027		-2,38	0,016	NS	NS
Glycolyse et néoglucogenèse	0,027			NS	NS	NS
Métabolisme des peptides	0,027		-4,074	0,0075	NS	NS
Métabolisme du tryptophane	0,032			NS	NS	NS
Métabolisme de la phénylalanine	0,032		-1,41	0,046	NS	NS
Oxydation des acides gras	0,032			NS	NS	NS
Voie des pentoses phosphates	0,033			NS	NS	NS
Biosynthèse des aminoacyl-ARNt	0,033		-3,85	0,0088	NS	NS
Dégradation des protéines	0,034		-3,85	0,012	NS	NS
Métabolisme des médicaments	0,036			NS	NS	NS
Métabolisme des leucotriènes	0,037			NS	NS	NS
Métabolisme du folate	0,038			NS	NS	NS
Cycle des acides tricarboxyliques et métabolisme du glyoxylate et des dicarboxylates	0,042			NS	NS	NS
Métabolisme du nicotinate et de la nicotinamide	0,045			NS	NS	NS
Métabolisme des porphyrines	0,045			NS	NS	NS
Métabolisme de la vitamine B6	0,049		-0,77	0,048	NS	NS
Métabolisme des purines	NS		-4,48	0,0059	NS	NS

f) Relation effet-dose

Nous avons étudié la relation effet-dose entre la plombémie et les voies métaboliques. Pour cela, nous avons étudié les voies métaboliques significatives avec la méthode du Z-score. Nous avons observé une diminution du glutathion avec l'augmentation de la plombémie (figures III.E.3.6 et III.E.3.7). Les résultats des autres voies métaboliques suivent la même tendance et sont présentées en **annexe 6**.

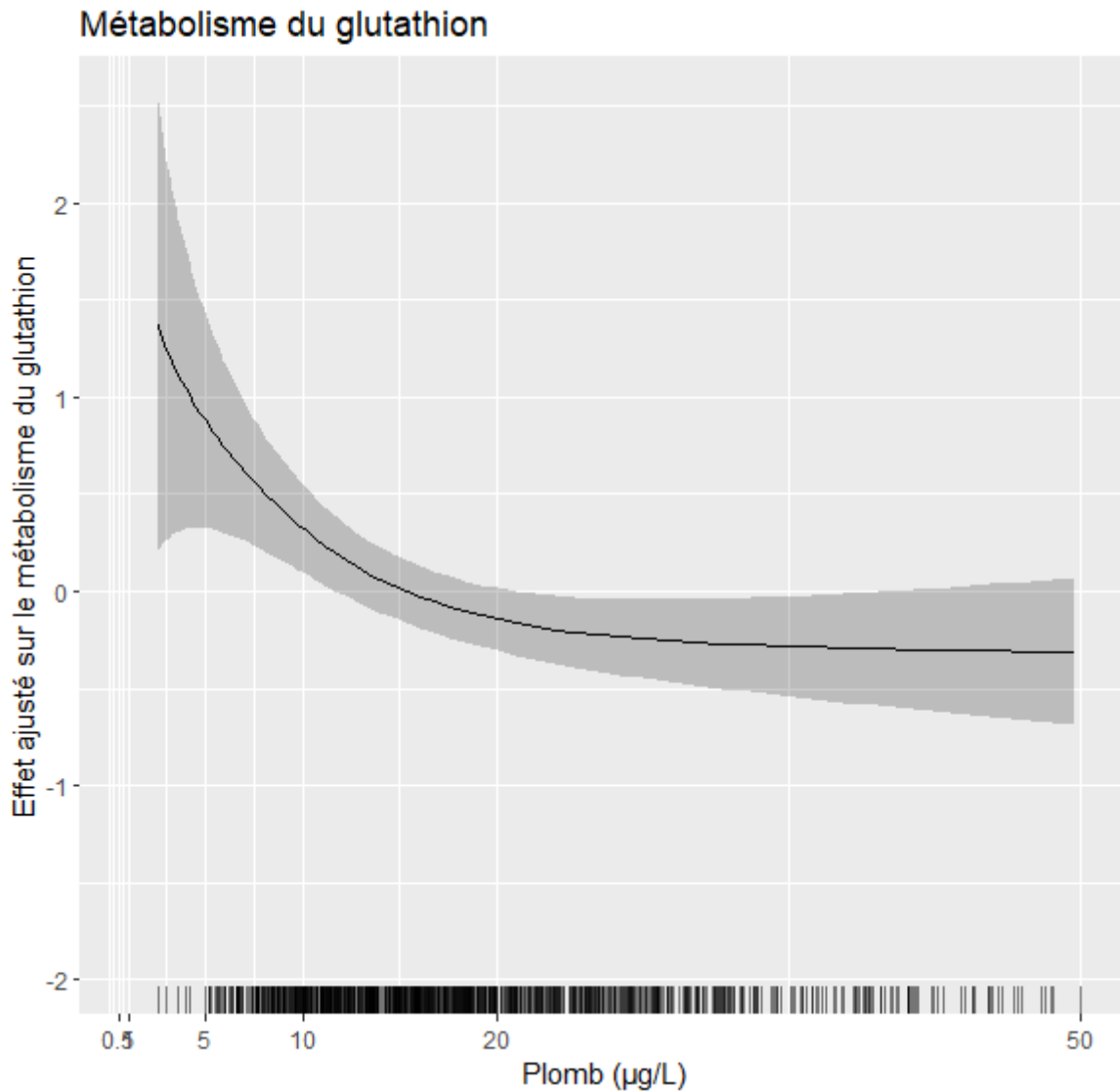


Figure III.E.3.6 : Courbe de la relation effet-dose entre la plombémie et la voie métabolique du glutathion

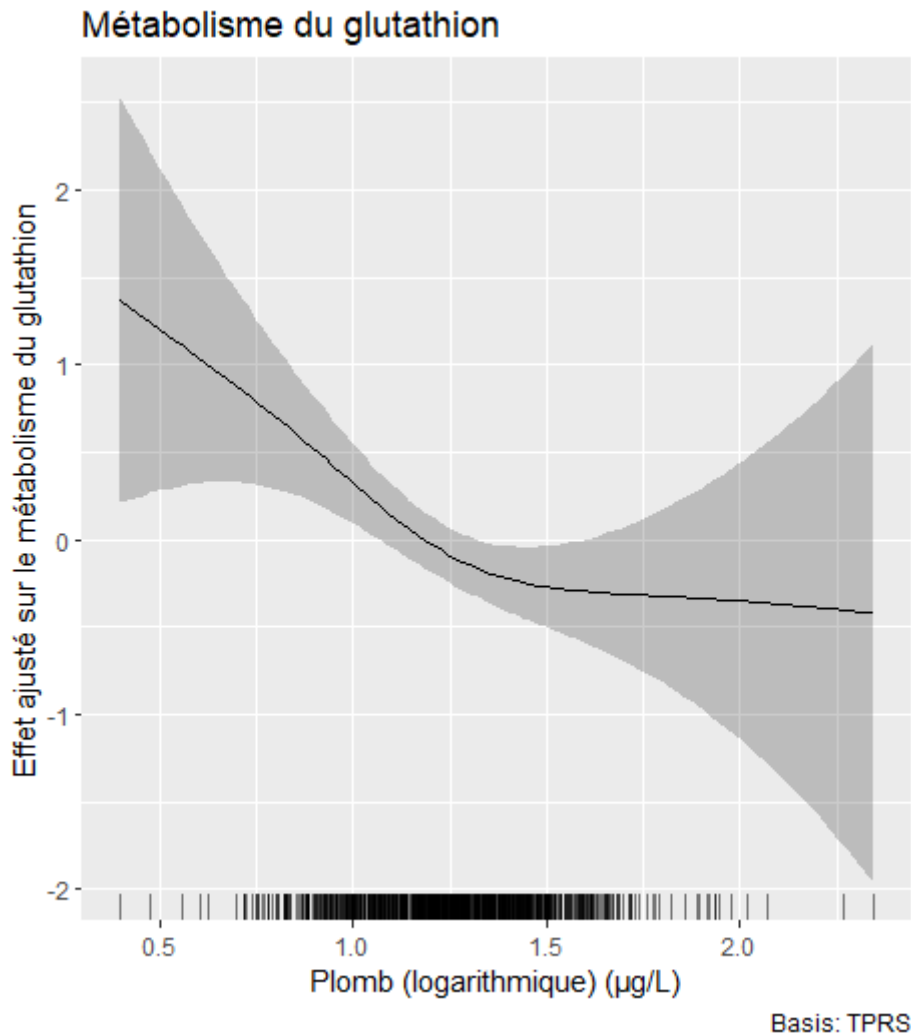


Figure III.E.3.7 : Courbe de la relation effet-dose entre la plombémie (logarithmique) et la voie métabolique du glutathion

4. Discussion de cette étude

a) Comparaison des 4 méthodes pour étudier les voies métaboliques

Nous avons 50 voies métaboliques associées à l'exposition au plomb par la méthode du rapport de vraisemblance, 18 par la méthode Z score, 1 par l'ORA et aucune par MSEA.

Les méthodes d'enrichissement de voies métaboliques (ORA et MSEA) n'ont pas ou peu identifié de voies métaboliques. Comme discuté dans la revue systématique, une des limites de l'ORA est qu'elle nécessite de filtrer les métabolites en fonction de la valeur de p et du *fold change* (exponentielle du coefficient d'association entre l'exposition et le métabolite). Des effets de seuil peuvent affecter les métabolites dont la valeur de p est proche du seuil de significativité, et leur inclusion ou exclusion de l'analyse d'enrichissement peut être fortement influencée par des facteurs aléatoires.

Avec la MSEA aucune voie métabolique n'était significative. Cela peut s'expliquer par le fait qu'un grand nombre de voies métaboliques sont affectées et, par conséquent, entrent en compétition les unes avec les autres. Par définition, toutes les voies ne peuvent pas être surreprésentées, ce qui conduit à un résultat non significatif avec cette méthode.

La méthode basée sur les scores Z semble présenter une plus grande puissance en raison d'un nombre plus faible de degrés de liberté. Toutefois, l'une de ses limites est que si certains métabolites au sein d'une même voie sont augmentés tandis que d'autres sont diminués, le score Z global peut être nul, rendant ainsi l'analyse de la voie non informative.

Le test du rapport de vraisemblance a montré davantage de voies métaboliques significatives. Cette méthode semble avoir plus de puissance.

Ce test implique un grand nombre de degrés de liberté, ce qui le rend plus adapté lorsque le nombre de variables n'est pas élevé, comme avec les voies métaboliques.

b) Comparaison avec la littérature

Les voies métaboliques retrouvées associées au plomb varient d'une étude à l'autre. Par exemple, une étude menée chez 99 femmes enceintes à Mexico (274), a mesuré les métabolites dans le sang, et fait une ORA pour rechercher les voies métaboliques associées au plomb sanguin. Ces voies étaient principalement des voies portant sur des protéines, c'est à dire des voies qui peuvent être impactées par de nombreuses perturbations métaboliques et qui peuvent être également un effet indirect des perturbations induites par le plomb. En effet, les voies métaboliques des protéines peuvent être impactées car reliée au métabolisme global. Dans leur étude en population générale incluant 399 hommes, Kelly *et al.*, ont mesuré les métabolites dans le plasma et ont retrouvé des associations entre la plombémie et la voie métabolique correspondant au métabolisme des glycérophospholipides et celle correspondant au métabolisme des sphingolipides (278). La recherche des voies métaboliques associées a été faite par une ORA. Nous avons retrouvé ces deux voies métaboliques par la méthode du rapport de vraisemblance mais pas avec les autres méthodes. Cependant, Kelly *et al.*, ont étudié les métabolites sanguins et non urinaires comme dans notre étude, ce qui peut expliquer les différences de voies métaboliques retrouvées. De plus, les différences d'effectifs peuvent expliquer les variations dans les voies associées selon les études.

Les concentrations sanguines en plomb observées dans notre étude n'étaient pas supérieures à celles mesurées dans la population française de l'étude nationale Esteban (la médiane dans notre étude était de 17,7 µg/L, contre 22,03 µg/L dans la tranche d'âge 45-59 ans et 27,57 µg/L dans la tranche d'âge 60-74 ans de l'étude Esteban) (39).

c) Physiopathologie

Pour la discussion sur l'investigation physiopathologique, nous nous concentrons sur les voies métaboliques retrouvées avec le rapport de vraisemblance et le Z-score. Ces voies métaboliques sont retrouvées avec au moins 2 méthodes statistiques différentes, ce qui suggère une robustesse des analyses et pourrait limiter le risque de faux positif, et permettrait d'identifier les voies les plus impactées.

Voies métaboliques liées au groupement thiol

Le plomb a une affinité pour le groupement thiol (-SH) présent sur de nombreuses molécules comme le glutathion, la cystéine et l'homocystéine. L'impact du plomb sur les thiols entraîne un stress oxydant par la production d'espèce réactives de l'oxygène (282).

Plusieurs voies métaboliques (métabolisme de la cystéine et de la méthionine, métabolisme du glutathion, métabolisme de l'histidine, métabolisme du soufre, métabolisme de la vitamine B6) significativement associées à la plombémie interviennent dans le métabolisme de l'homocystéine, elle-même associée aux facteurs de risque cardiovasculaires. En effet, l'homocystéine, n'est pas un marqueur utilisé en clinique mais est associé à des maladies cardiovasculaires dans des études épidémiologiques sans que le lien causal ne soit clairement établi (283). Même s'il n'y a pas de lien causal clairement établi, l'homocystéine pourrait être un marqueur d'altération de processus physiopathologiques associés au risque de maladies cardiovasculaires. L'association entre l'exposition au plomb et ce métabolisme est cohérente avec l'augmentation du risque cardiovasculaire associé à l'exposition au plomb (284). De plus, l'exposition au plomb est associée à un taux élevé d'homocystéine (285).

L'homocystéine subit deux réactions principales. Une reméthylation où l'homocystéine est transformée en méthionine, acide aminé sulfuré, par l'intermédiaire de la vitamine B12 et des folates, en particulier le 5-méthyl tétrahydrofolate (THF), produit par le métabolisme de la sérine-glycine. La deuxième réaction est la transsulfuration où l'homocystéine est transformée en cystéine, via la sérine et la vitamine B6. La cystéine est un précurseur du glutathion, composée de 3 acides aminés : le glutamate, la cystéine et la glycine. Par ailleurs, l'histidine, participe au cycle des folates par son catabolisme en acide formiminoglutamique (FIGLU). Et la thréonine peut être catabolisée en glycine. Par ailleurs, le plomb a une affinité pour le radical thiol (comportant un atome de soufre). Par ailleurs, le plomb est impliqué dans le stress oxydant. Dans la littérature il a été observé une augmentation des espèces réactives de l'oxygène avec l'exposition au plomb entraînant une augmentation du calcium intracellulaire conduisant à l'apoptose des cellules neuronales (286).

Dans la littérature, des associations entre une exposition au plomb et des perturbations de ce groupe de voies métaboliques ont été retrouvées. Une étude s'est intéressée à des travailleurs exposés au plomb et a retrouvé une diminution de la méthionine et une augmentation de l'histidine (287). Une autre étude a regardé l'effet de l'exposition au plomb sur l'appareil reproducteur chez des rats et a montré qu'une supplémentation en histidine réduisait le stress oxydatif et réduisait le dysfonctionnement mitochondrial induit par le plomb dans le système reproducteur des rats (288).

Dans une étude portant chez des adultes aux Etats-Unis et s'intéressant aux effets du plomb sur les maladies cardiovasculaires, il a été montré qu'un taux sanguin de plomb plus élevé était associé à un taux de vitamine B6 plus faible. De plus, dans leur groupe carencé en vitamine B6, le plomb était associé aux maladies cardiovasculaires (289).

Le glutathion était augmenté dans les érythrocytes de travailleurs exposés (290) et était augmenté dans les érythrocytes, les reins et le foie chez les rats (291).

Certaines voies métaboliques identifiées sont liées au stress oxydant : métabolisme du glutathion et métabolisme de la thiamine. Comme expliqué au-dessus, le glutathion présente

un groupement thiol dont le plomb a une affinité. Le glutathion joue un rôle clé dans la régulation du stress oxydant en neutralisant les espèces réactives de l'oxygène (ERO) (43). La thiamine, ou vitamine B1, a une action antioxydante également (262).

Groupe des voies métabolique lié au glucose

Des voies métaboliques liées au glucose étaient également perturbées dans notre étude : la voie du métabolisme des sucres aminés et des nucléotides-oses, la voie du métabolisme des propanoate, la voie du métabolisme des O-glycanes et la voie du métabolisme de la thiamine. Les sucres aminés et les nucléotides-oses sont indispensables à la construction des glycoprotéines et ont un rôle majeur dans la glycosylation (mécanisme qui ajoute des sucres (glycanes) sur une protéine ou un lipide) (292). Le propanoate est une source de glucose via le cycle de Krebs et la néoglucogenèse et a un rôle de régulation de la glycémie (293). Enfin, la thiamine est indispensable au métabolisme du glucose (262).

Dans la littérature, une étude a montré qu'une exposition chronique à de faibles doses de plomb peut entraîner une augmentation de la production hépatique de glucose via la surexpression des enzymes clés de la néoglucogenèse (294). En outre, une exposition au plomb augmente l'activité des enzymes hépatiques de la néoglucogenèse entraînant une intolérance au glucose (295).

Autres voies métaboliques

D'autres groupes de voies métaboliques ont été identifiés.

Tout d'abord, des voies métaboliques liées aux protéines ont été retrouvées : métabolisme des peptides, dégradation des protéines, biosynthèse de la phénylalanine, de la tyrosine et du tryptophane, métabolisme de la phénylalanine et biosynthèse des aminoacyl-ARNt.

La voie métabolique de la phénylalanine, de la tyrosine et du tryptophane semble liée à la phénylalanine hydroxylase (PAH). Une hypothèse pourrait être que le plomb interfère avec la synthèse des protéines de manière aspécifique par l'intermédiaire des autres voies métaboliques perturbées. Une autre hypothèse serait que l'association entre l'exposition au plomb et les protéines n'est pas encore connue et serait à explorer.

Ensuite, des voies métaboliques liées à l'ARN ou l'ADN ont été retrouvées et concernent le métabolisme des nucléotides et le métabolisme de la pyrimidine, base azotée des nucléotides. Enfin, la voie de la biosynthèse du squelette des terpénoïdes était associée à la plombémie. Les terpénoïdes produisent des stérols. Cette voie peut affecter les lipides.

d) Relation effet-dose

Les courbes évaluant la relation effet-dose ont montré l'absence de valeur seuil en-dessous de laquelle aucun effet ne serait observé. Pour la voie métabolique du glutathion, par exemple, la relation est compatible avec une association linéaire avec le logarithme de l'exposition au plomb. Cette relation linéaire avec le logarithme suggère une pente plus importante pour les expositions plus faibles au plomb (en valeur non transformée) qu'aux expositions élevées. Ceci suggère un possible effet de saturation relative des impacts du plomb sur les voies

métaboliques pour les fortes doses. Ces courbes suggèrent, des effets sur le métabolisme dès de faibles doses. Ces résultats soulèvent la question des normes chez les adultes pour l'exposition au plomb.

Forces et limites

Cette étude comporte plusieurs forces. Tout d'abord, les fumeurs n'étaient pas inclus dans notre étude, de sorte que l'effet métabolique observé ne peut pas être attribué au tabac en tant que facteur de confusion. La mesure de l'exposition au plomb par dosage sanguin plutôt que dans les urines est recommandée pour évaluer une potentielle intoxication au plomb. De plus, nous avons utilisé différentes méthodologies statistiques dont le test du rapport de vraisemblance qui est un test simple à utiliser et qui ne nécessite pas d'hypothèses.

Mais elle comporte aussi certaines limites. L'exclusion des fumeurs limite la généralisabilité des résultats. La mesure du plomb dans le sang peut à la fois rendre compte d'une exposition récente ou ancienne. De plus, le grand nombre de voies métaboliques retrouvées, avec notamment le test du rapport de vraisemblance, rend difficile la synthèse des résultats. Enfin, il est très difficile de quantifier l'impact réel de ces perturbations métaboliques sur des événements de santé, comme des événements cardiovasculaires.

5. Conclusion de cette étude

La population de cette étude est peu exposée au plomb. De nombreuses associations avec les métabolites ont été retrouvées, même à ces faibles doses d'exposition. De nombreuses voies métaboliques étaient également impactées par le plomb. Parmi ces perturbations métaboliques, un nombre important peut être expliqué par l'affinité du plomb pour les thiols et par son impact sur le métabolisme du glutathion, impliqué dans la régulation du stress oxydant, ce qui renforce la plausibilité biologique des résultats.

L'étude de la relation effet-dose a montré l'absence de valeur seuil, suggérant des associations à faibles doses. Ces résultats suggèrent que des perturbations métaboliques pourraient exister à des seuils très inférieurs aux normes d'exposition. L'impact à long terme sur la santé reste inconnu ; néanmoins cela suggère un intérêt à continuer à réduire l'exposition au plomb.

IV. DISCUSSION GENERALE

A. Etudes de variables intermédiaires : facteurs de risque connus et biomarqueurs

Cette thèse explore les liens entre des expositions environnementales et la santé en utilisant différentes approches.

1. Etude d'intermédiaires : facteurs de risque

Tout d'abord, nous avons étudié ces liens par l'intermédiaire de facteurs de risque cardiovasculaire connus. Nous avons retrouvé une association négative entre l'exposition à la pollution atmosphérique et les taux sanguins de HDL. Ces résultats viennent compléter les travaux antérieurs réalisés dans ELISABET et qui avaient retrouvé une augmentation de l'HbA1c avec l'exposition à la pollution atmosphérique (48) et au cadmium (49).

Cette approche méthodologique ancienne est relativement simple et toujours pertinente.

Elle se base sur un lien connu et quantifié avec le risque réel de pathologie. Elle peut se faire en étude transversale, sur des effectifs moins importants que des cohortes longitudinales, ce qui en fait une approche pragmatique pour explorer rapidement des liens potentiels entre l'environnement et des marqueurs intermédiaires, sans nécessiter un suivi prolongé. Cependant cette approche reste essentiellement limitée au risque cardiovasculaire où ces facteurs de risque bien connus et dépistés dans le cadre du soin courant existent.

2. Etude d'intermédiaires : biomarqueurs

Nous avons également étudié des biomarqueurs. L'utilisation de biomarqueurs témoins d'un processus intermédiaire entre exposition et pathologie ou facteur de risque est une approche potentiellement plus sensible aux faibles expositions car ces biomarqueurs pourraient être modifiés de manière précoce.

Dans ELISABET, certaines associations étaient retrouvées, mais, pour la majorité des biomarqueurs aucune association significative n'était retrouvée malgré un échantillon d'environ 1 000 personnes.

En effet, dans ce travail de thèse, des associations positives ont été retrouvées entre l'exposition à court terme à la pollution atmosphérique et le TNF α et la CRPus (pour ces analyses sur la CRPus, l'effectif était plus grand (environ 3 000 personnes)).

La CRPus est un marqueur assez ancien et connu de l'inflammation de bas-grade (296).

Cependant les associations avec les 6 autres interleukines étaient non significatives (54).

Dans les autres travaux sur ELISABET, une étude avait des associations significatives avec 5 micro-ARN sur les 10 testés (52) tandis qu'une autre n'avait aucune association significative avec les 4 marqueurs du stress oxydant testés (8-hydroxydeoxyguanosine (8-OHdG), 4-hydroxynonanal (4-HNE), fluorescent oxidation, products (FOPs), Club cell secretory protein (CC16)) (53).

Les résultats non significatifs sur de nombreux biomarqueurs peuvent s'expliquer par une absence réelle d'association, ou un choix non pertinent des biomarqueurs. D'autres facteurs peuvent limiter la puissance de l'étude de ces associations. D'une part, l'erreur de mesure qui peut être importante sur des dosages réalisés dans un cadre de recherche pour lesquels les

standards techniques et l'expérience n'est pas la même que pour les marqueurs utilisés en pratique clinique. D'autre part, la variabilité temporelle de certains biomarqueurs pour un même individu peut être importante. L'analyse de ces biomarqueurs nécessiterait de réaliser des études longitudinales avec des mesures répétées plutôt que dans une étude transversale comme ELISABET.

B. Métabolomique

1. Fishing expedition

L'approche métabolomique permet d'avoir une démarche sans a priori. Elle permet d'avoir beaucoup de données que l'on peut croiser avec des données environnementales et éventuellement avec des données de physiologie. L'avantage de la métabolomique est qu'elle permet de disposer de nombreuses données. Mais elle présente le risque de multiplier les analyses sans objectifs clairement déterminés : « fishing expédition », (58). Dans notre revue de la littérature, nous avons constaté que les hypothèses de recherche n'étaient pas toujours faciles à identifier, qu'il n'y avait pas toujours d'alignement entre les questions de recherche, les hypothèses statistiques et les analyses statistiques, conduisant à une difficulté d'identifier l'apport des résultats à la connaissance sur le sujet étudié.

2. Importance de poser les hypothèses

Une approche sans a priori comme celle de la métabolomique n'exclut pas de poser les hypothèses H0 et H1. La génomique, est aussi une approche sans a priori. Dans les Genome-Wide Association Studies (GWAS), les associations entre gène et phénotype sont étudiées selon des stratégies faisant l'objet d'un consensus. Bien que les GWAS évaluent des centaines de milliers de gènes indistinctement et leur association avec un phénotypes cela ne signifie pas qu'il n'y a pas d'hypothèse posée. L'hypothèse testée est qu'il existe au moins une association entre un des gènes et le phénotype (297).

Le contexte de la métabolomique est différent. Mais l'exemple montre qu'il est possible d'arriver à une approche consensuelle et robuste. Si la démarche générale de la communauté scientifique de standardisation des méthodologies des GWAS est probablement un exemple à suivre, cela ne signifie pas que les méthodes spécifiques utilisées en génétique sont adaptées à la métabolomique. Les méthodes d'analyse de la métabolomique s'inspirent souvent de la génétique. C'est le cas, par exemple, de la MSEA. Nous avons vu dans ce travail que cette méthode n'était pas nécessairement la plus performante pour l'analyse des voies métaboliques en métabolomique, et qu'il pouvait être intéressant de développer des méthodologies alternatives, comme le rapport de vraisemblance.

Notre revue de la littérature met en évidence l'importance de poser correctement et précisément la question scientifique de recherche, d'en déduire les hypothèses H0 et H1, pour pouvoir ainsi mieux adapter les méthodes statistiques.

Un exemple de l'imprécision qui peut exister dans la question scientifique est l'usage fréquent de termes comme « profil » ou « signature » dans les titres ou les objectifs dont nous avons vu qu'ils pouvaient recouvrir des sens différents selon les auteurs, sans être clairement définis.

Nous faisons une proposition permettant d'aider à orienter la stratégie d'analyse en fonction de la question scientifique.

Si la question de recherche est : « Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ? », alors une hypothèse H1 pertinente est : « il existe une association entre au moins une voie métabolique et une exposition environnementale ». L'outcome statistique attendu est alors une analyse par voie métabolique à l'aide de méthodes statistiques adaptées comme le Z-score ou le rapport de vraisemblance. Il est également possible de réaliser une analyse par métabolites individuels avec une correction pour les tests multiples par FDR en cas des régressions.

Si la question de recherche est : « Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ? », les hypothèses H1 attendues sont « au moins un métabolite médie la relation entre l'exposition et le phénotype », « au moins une voie métabolique médie la relation entre l'exposition et le phénotype » ou « au moins un cluster médie la relation entre l'exposition et le phénotype ». L'outcome statistique est pertinent s'il concerne l'association entre l'exposition et la variable de médiation, puis l'association entre la variable de médiation et le phénotype. Une condition importante est que la variable de médiation doit être indépendante de l'exposition et du phénotype. Sinon il faut un échantillon de construction et de validation lors de la réalisation des analyses statistiques dont la méthode choisie doit inclure le calcul de la part médiée.

Enfin, si la question de recherche est : « L'exposition influence-t-elle le métabolisme ? », alors l'hypothèse H1 est de type : « il existe une association entre au moins une composante ou cluster et une exposition ». L'outcome statistique attendu est une composante ou cluster et les méthodes statistiques sont basées sur des analyses de type multidimensionnel : ACP, analyse de cluster ou encore PLS. Il convient de noter que pour l'utilisation de PLS, l'hypothèse H1 est difficile à formuler.

Le tableau IV.B.2.1 ci-dessous résume notre proposition de démarche.

Tableau IV.B.2.1 : Stratégie d'analyse en métabolomique

Question scientifique	Hypothèses H1 possibles	Outcome statistique attendu	Méthodes statistiques appropriées
Quels processus métaboliques ou quelles voies métaboliques spécifiques sont influencés par l'exposition ?	Il existe une association entre au moins un métabolite et une exposition environnementale Il existe une association entre au moins une voie métabolique et une exposition environnementale (meilleure hypothèse)	Analyse par voies métaboliques +/- une analyse par métabolites individuels	Analyse par voie métabolique : Z score ou rapport de vraisemblance Analyses par métabolites : régression (avec correction par FDR) ou Lasso ou elastic net (échantillon de validation pour Lasso et elastic net)
Les métabolites sont-ils des médiateurs de l'association entre l'exposition et le phénotype ?	Au moins un métabolite médie la relation entre l'exposition et le phénotype Au moins une voie métabolique médie la relation entre l'exposition et le phénotype Au moins un score métabolique médie la relation entre l'exposition et le phénotype	Association entre l'exposition et la variable de médiation, association entre la variable de médiation et le phénotype Variable de médiation (métabolites, voie métaboliques, score, composantes) dont on calcule la part médiée La variable de médiation doit être indépendante de l'exposition et du phénotype	Analyses incluant le calcul de la proportion médiée Si la variable de médiation n'est pas indépendante de l'exposition et du phénotype, il faut un échantillon de construction et de validation
L'exposition influence-t-elle le métabolisme ?	Il existe une association entre au moins une composante ou cluster et une exposition	Composante ou cluster	Analyses de type multidimensionnel : ACP, analyse de cluster, PLS* Si PLS : échantillon de construction et de validation

* Si PLS alors l'hypothèse H1 est difficile à formuler

Les conclusions des articles inclus dans notre revue de la littérature étaient souvent décevantes. En effet, il est fréquent que les auteurs énumèrent leurs résultats, sans conclusions claires quant à l'apport de ceux-ci. Nous pouvons néanmoins supposer qu'avec des hypothèses de travail mieux posées et une méthodologie mieux cadrée, la métabolomique serait une approche robuste et son apport plus facile à mettre en évidence.

En 2007, Miller et al., (58) alertaient déjà sur le risque de déception par rapport à la métabolomique, en plein essor à cette époque (« A real threat to this emerging field is that, after all the enthusiasm, there is disillusionment with what metabolomics can and cannot realistically deliver. If the approach is "oversold", a general skepticism can result. As more metabolomic data is collected, there is the possibility that what is detected cannot be interpreted with respect to risk to human health and the environment »).

Ce risque pourrait provenir du contraste entre la performance croissante des méthodes d'analyse biologique et le nombre croissant de données disponibles d'un côté, et, de l'autre une difficulté à construire une méthodologie d'analyse robuste permettant d'extraire de façon claire les informations pertinentes de ces données. Nous avons vu dans notre revue de la littérature que ce risque semble présent dans nombreux travaux. A l'inverse l'exemple de nos analyses sur le plomb montre qu'en essayant de poser une stratégie d'analyse cohérente, on pouvait identifier des perturbations métaboliques que l'on pouvait expliquer par des mécanismes physiopathologiques connus et liés à l'exposition et arriver à une conclusion cohérente.

C. Etudes des liens entre expositions et métabolomique dans ELISABET

1. Choix des analyses sur l'ensemble des expositions

Nous nous sommes appuyés sur les propositions issues de notre revue de la littérature pour étudier les liens entre les expositions disponibles dans ELISABET et les données de métabolomique (provenant de Metalisabet). Nous avons fait le choix de présenter toutes nos analyses c'est-à-dire toutes les expositions environnementales testées, afin d'éviter un biais lié à la sélection a posteriori des résultats.

2. Discussion des résultats obtenus

Pour nos résultats, la majorité des associations sont non significatives. Cependant, l'absence de significativité ne signifie pas l'absence d'un impact sur le métabolisme.

En effet, la taille d'effet d'une association statistique est fortement impactée par la qualité de la mesure. L'erreur de mesure non différentielle qui est souvent importante dans l'évaluation des expositions environnementales tend à diminuer la taille d'effet observée et la puissance statistiques des études. Par exemple, l'exposition à la pollution atmosphérique est mesurée dans ELISABET de manière indirecte. Pour l'exposition à long terme elle correspond à une modélisation statistique à l'adresse du domicile ; accumulant les erreurs liées aux imprécisions de la modélisation et au fait que l'exposition au domicile n'est qu'une partie de l'exposition de personnes qui se déplacent dans des environnements intérieurs et extérieurs. Cela entraîne des imprécisions de mesure et probablement une sous-estimation de l'exposition des

participants. A l'inverse, la plombémie est une mesure directe individuelle qui est susceptible d'être beaucoup plus représentative de l'exposition réelle des individus. Il est donc difficile de comparer les tailles d'effet et les nombres de métabolites ou de voies métaboliques impactés ou d'établir une hiérarchie entre ces différentes expositions.

3. Exposome

Ce point illustre la difficulté d'étudier l'exposome. L'exposome, est défini par Rappaport et al., (298) comme l'ensemble de toutes les substances chimiques biologiquement actives présentes dans le sang d'une personne - issues à la fois de sources exogènes et de processus endogènes. L'étude de l'exposome ne peut pas se calquer sur celle du génome. Par exemple, une technique unique permet d'étudier l'ensemble du génome avec une nature de variable et une qualité de mesure comparable pour chaque gène. A l'inverse l'étude de l'exposome est toujours très partielle et limitée par des natures d'exposition et de qualités mesures très variables. La quantification de l'impact de l'environnement et de chacune de ses composantes, des multi-expositions et des interactions est complexifiée par ces problèmes de mesure, même dans des cohortes où l'exposition serait renseignée à l'aide des meilleurs outils existants.

4. Exemple du plomb

Enfin, nous avons développé dans cette thèse l'exemple de l'exposition au plomb. L'intérêt est de réfléchir à l'interprétation des résultats obtenus. En effet, la « fishing expedition » soulève le problème de l'interprétation a posteriori des résultats (58).

L'exemple du plomb montre qu'il y a de nombreuses voies métaboliques impactées. Il est difficile de toutes les explorer. Nous nous sommes restreints à celles retrouvées par différentes méthodes. Nous avons supposé que ces voies étaient les plus importantes en termes d'impact sur le métabolisme mais cette hypothèse est potentiellement fautive.

Nous nous sommes heurtés à la difficulté d'identifier les voies métaboliques les plus impactées car même une voie éloignée du métabolisme initial peut être impactée. Il est très complexe de hiérarchiser toutes ces voies retrouvées.

Malgré cette difficulté, nous avons réussi à identifier au moins un mécanisme physiologique connu susceptible d'expliquer une part importante des perturbations observées. Ce mécanisme correspond à l'affinité du plomb pour le groupement thiol. Cet exemple montre la possibilité d'arriver à une interprétation des résultats cohérente avec la physiologie et d'améliorer la compréhension et la quantification des mécanismes en jeu grâce à la métabolomique. Néanmoins il est possible que nos conclusions soient alimentées par un biais de confirmation ou la reconstruction a posteriori d'une explication cohérente et logique à des phénomènes qui seraient en réalité aléatoires ou sous-tendus par des explications très différentes dans un mécanisme se rapprochant du phénomène de paréidolie.

D. Perspectives

Les perspectives à court terme sont :

- De publier les articles concernant la revue et l'exposition au plomb
- D'analyser plus en détail les voies métaboliques pour les expositions pour lesquelles nous avons des métabolites significatifs
- De publier les résultats non significatifs

Les perspectives à moyen terme sont de travailler sur l'identification de nouvelles méthodes statistiques permettant de tenir compte des liens a priori entre les métabolites et également de quantifier les zones du métabolisme les plus impactées afin que l'interprétation des mécanismes modifiés soit moins subjective.

En effet, nous avons identifié 2 limites importantes pour les méthodes actuelles :

- Lorsqu'il y a beaucoup de voies métaboliques impactées, il est difficile de trouver le ou les mécanismes initiaux les plus impactés par l'exposition qui seraient témoins de l'action initiale de l'exposition environnementale sur les modifications du métabolisme.
- Les analyses statistiques portent soit sur les métabolites, soit sur des dimensions ou clusters définis sans a priori. Les seuls tests statistiques qui tiennent compte des relations connues a priori entre métabolites sont ceux portant sur les voies métaboliques ; qui constituent une manière assez grossière d'aborder la complexité des réseaux métaboliques. Nous n'avons pas trouvé de méthode statistique prenant en compte l'entièreté du réseau métabolique dans toute sa complexité pour la construction des hypothèses et des tests statistiques. Une telle approche pourrait constituer une voie prometteuse pour augmenter la puissance des analyses statistiques et aboutir à des résultats plus facilement interprétables.

Nous avons identifié ce questionnement récemment, cette approche reste très spéculative.

Les réseaux lillois pluridisciplinaires comme ECRIN (Environnement Climat Recherche & Innovation) devraient nous permettre de nouer des collaborations avec des équipes d'informatiques comme le laboratoire CRISAL (Centre de Recherche en Informatique, Signal et Automatique de Lille) ou l'équipe METRICS (dont la thématique de travail est centrée sur l'évaluation des technologies de santé et des pratiques médicales) de l'université de Lille, pour tenter d'avancer sur cette approche.

Les perspectives à long terme seraient d'étudier les variations du métabolome dans des études longitudinales pour étudier les effets à court terme sur le métabolisme de certaines expositions et prendre en compte la variabilité biologique intra individuelle. Ce type d'étude inclurait à la fois des données répétées de mesure sur le métabolisme mais aussi une mesure longitudinale de l'exposition environnementale, par exemple, par l'utilisation de microcapteurs portables de mesure des PM.

La variabilité inter- et intra individuelle des métabolites peut entraîner des résultats faussement négatifs. Ainsi, même si la métabolomique est décrite comme très sensible, elle peut malgré tout ne pas identifier d'impact existant notamment dans les études transversales. Les mesures répétées dans une approche longitudinale pourraient réduire cette limite.

Nous souhaiterions des données individuelles répétées sur le métabolisme et des microcapteurs pour mieux estimer les expositions individuelles, afin d'explorer ces variations à court terme. Nous initiions des projets d'études longitudinales à Lille. Le recrutement du projet Mobicard va bientôt débuter (responsable Stephan Gabet, IMPECS, investigateur principal Luc Dauchet), celui-ci a pour objectif de caractériser en détail l'exposition individuelle et simultanée des habitants de l'agglomération lilloise à plusieurs polluants contribuant à l'exposome, ainsi que leur variabilité sur une semaine ; de mettre en évidence les effets de la mobilité et des activités, physiques ou non, sur les niveaux d'exposition et les paramètres

cardiovasculaires ; et de constituer une bibliothèque (sang et urine) pour étudier à terme plusieurs biomarqueurs d'exposition et d'effet (299).

V. CONCLUSION

L'étude de facteurs intermédiaires entre l'exposition environnementale et la santé, tels que les facteurs de risque, les biomarqueurs et la métabolomique, permettrait de mieux comprendre les mécanismes physiopathologiques expliquant ces relations et d'identifier des liens entre environnement et santé lorsque la taille d'effet est trop faible ou que le suivi est insuffisant pour mettre en évidence un impact sur les événements de santé. Cette approche pourrait donc être intéressante pour identifier et quantifier indirectement l'effet de faibles doses d'exposition en population générale, et d'identifier les expositions les plus suspectes d'avoir un impact afin d'orienter les priorités de santé publique en santé environnement.

Nos travaux ont permis de mettre en évidence une certaine efficacité de cette approche. Les associations avec les facteurs de risque permettent une évaluation indirecte du risque cardiovasculaire. L'utilisation de biomarqueurs ou de la métabolomique peut être une approche sensible comme l'ont montré les perturbations métaboliques très significatives associées à la plombémie malgré des niveaux d'exposition relativement faibles.

Cependant, notre travail a aussi mis en évidence des difficultés méthodologiques dans ces approches. Elles se heurtent en particulier à la difficulté de mesurer l'exposition avec précision et à la variabilité des marqueurs biologiques. De plus, notre revue de la littérature a montré une absence de standardisation des approches statistiques dans l'utilisation de la métabolomique dans ce cadre. Cela limite l'identification de l'apport de ces travaux à la connaissance scientifique. Nous avons proposé des pistes permettant de mieux adapter la méthodologie aux questions scientifiques posées. Les voies d'amélioration dans cette démarche passeraient par une amélioration de la qualité de la donnée d'exposition et par l'élaboration de consensus collectifs sur les stratégies d'analyses statistiques.

VI. REFERENCES

1. Smith KR, Corvalán CF, Kjellström T. How much global ill health is attributable to environmental factors? *Epidemiology*. sept 1999;10(5):573-84. PubMed PMID: 10468437.
2. ResearchGate [Internet]. [cité 11 oct 2025]. (PDF) Preventing disease through healthy environments: Towards an estimate of the environmental burden of disease. Disponible sur: https://www.researchgate.net/publication/237346974_Preventing_disease_through_healthy_environments_Towards_an_estimate_of_the_environmental_burden_of_disease
3. Collège Universitaire des Enseignants de Santé Publique. Item : 367 Impact de l'environnement sur la santé. Dauchet L, Viel Jf, Delva F. In.
4. Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks [Internet]. [cité 25 sept 2025]. Disponible sur: <https://www.who.int/publications/i/item/9789241565196>
5. Cartes de trafics annuels - DREAL HAUTS-DE-FRANCE [Internet]. [cité 28 déc 2021]. Disponible sur: <https://www.hauts-de-france.developpement-durable.gouv.fr/?Cartes-de-trafic-annuels>
6. https://www.hauts-de-france.developpement-durable.gouv.fr/IMG/pdf/recensement_circulation_region_nord_pas_de_calais_2012_trafic_tous_vehicules.pdf.
7. Dauchet L, Hulo S, Cherot-Kornobis N, Matran R, Amouyel P, Edmé JL, et al. Short-term exposure to air pollution: Associations with lung function and inflammatory markers in non-smoking, healthy adults. *Environ Int*. déc 2018;121(Pt 1):610-9. doi:10.1016/j.envint.2018.09.036 PubMed PMID: 30312964.
8. Air pollution [Internet]. [cité 13 sept 2025]. Disponible sur: <https://www.who.int/health-topics/air-pollution>
9. Pascal M, de Crouy Chanel P, Wagner V, Corso M, Tillier C, Bentayeb M, et al. The mortality impacts of fine particles in France. *Sci Total Environ*. 15 nov 2016;571:416-25. doi:10.1016/j.scitotenv.2016.06.213 PubMed PMID: 27453142.
10. Adélaïde L. IMPACT DE LA POLLUTION DE L'AIR AMBIANT SUR LA MORTALITÉ EN FRANCE MÉTROPOLITAINE : RÉDUCTION EN LIEN AVEC LE CONFINEMENT DU PRINTEMPS 2020 ET IMPACT À LONG TERME POUR LA PÉRIODE 2016-2019 / IMPACT OF AIR POLLUTION ON MORTALITY IN METROPOLITAN FRANCE: REDUCTION RELATED TO THE SPRING 2020 LOCKDOWN AND LONG-TERM IMPACT FOR 2016-2019. 11.
11. Pascal M, de Crouy Chanel P, Wagner V, Corso M, Tillier C, Bentayeb M, et al. The mortality impacts of fine particles in France. *Sci Total Environ*. 15 nov 2016;571:416-25. doi:10.1016/j.scitotenv.2016.06.213 PubMed PMID: 27453142.

12. Bono R, Romanazzi V, Bellisario V, Tassinari R, Trucco G, Urbino A, et al. Air pollution, aeroallergens and admissions to pediatric emergency room for respiratory reasons in Turin, northwestern Italy. *BMC Public Health*. 5 août 2016;16(1):722. doi:10.1186/s12889-016-3376-3
13. Madaniyazi L, Xerxes S. Outdoor air pollution and the onset and exacerbation of asthma. *Chronic Dis Transl Med*. juin 2021;7(2):100-6. doi:10.1016/j.cdtm.2021.04.003 PubMed PMID: 34136769; PubMed Central PMCID: PMC8180519.
14. Guarnieri M, Balmes JR. Outdoor air pollution and asthma. *Lancet*. 3 mai 2014;383(9928):1581-92. doi:10.1016/S0140-6736(14)60617-6 PubMed PMID: 24792855; PubMed Central PMCID: PMC4465283.
15. DeVries R, Kriebel D, Sama S. Outdoor Air Pollution and COPD-Related Emergency Department Visits, Hospital Admissions, and Mortality: A Meta-Analysis. *COPD*. févr 2017;14(1):113-21. doi:10.1080/15412555.2016.1216956 PubMed PMID: 27564008; PubMed Central PMCID: PMC8994423.
16. Solimini AG, Renzi M. Association between Air Pollution and Emergency Room Visits for Atrial Fibrillation. *Int J Environ Res Public Health*. 20 juin 2017;14(6):661. doi:10.3390/ijerph14060661 PubMed PMID: 28632149; PubMed Central PMCID: PMC5486347.
17. Braithwaite I, Zhang S, Kirkbride JB, Osborn DPJ, Hayes JF. Air Pollution (Particulate Matter) Exposure and Associations with Depression, Anxiety, Bipolar, Psychosis and Suicide Risk: A Systematic Review and Meta-Analysis. *Environ Health Perspect*. déc 2019;127(12):126002. doi:10.1289/EHP4595 PubMed PMID: 31850801; PubMed Central PMCID: PMC6957283.
18. Huang Y, Zhu M, Ji M, Fan J, Xie J, Wei X, et al. Air Pollution, Genetic Factors, and the Risk of Lung Cancer: A Prospective Study in the UK Biobank. *Am J Respir Crit Care Med*. 1 oct 2021;204(7):817-25. doi:10.1164/rccm.202011-4063OC PubMed PMID: 34252012.
19. Li YC, Chiou JY, Lin CL, Wei JCC, Yeh MH. The association between air pollution level and breast cancer risk in Taiwan. *Medicine (Baltimore)*. 14 mai 2021;100(19):e25637. doi:10.1097/MD.00000000000025637 PubMed PMID: 34106591; PubMed Central PMCID: PMC8133213.
20. So R, Chen J, Mehta AJ, Liu S, Strak M, Wolf K, et al. Long-term exposure to air pollution and liver cancer incidence in six European cohorts. *Int J Cancer*. 1 déc 2021;149(11):1887-97. doi:10.1002/ijc.33743 PubMed PMID: 34278567.
21. Hansen AB, Ravnskjaer L, Loft S, Andersen KK, Brauner EV, Baastrup R, et al. Long-term exposure to fine particulate matter and incidence of diabetes in the Danish Nurse Cohort. *Environ Int*. mai 2016;91:243-50. doi:10.1016/j.envint.2016.02.036 PubMed PMID: 26989812.
22. Riant M, Meirhaeghe A, Giovannelli J, Occelli F, Havet A, Cuny D, et al. Associations between long-term exposure to air pollution, glycosylated hemoglobin, fasting blood

- glucose and diabetes mellitus in northern France. *Environ Int.* nov 2018;120:121-9. doi:10.1016/j.envint.2018.07.034 PubMed PMID: 30077944.
23. Yang BY, Bloom MS, Markevych I, Qian ZM, Vaughn MG, Cummings-Vaughn LA, et al. Exposure to ambient air pollution and blood lipids in adults: The 33 Communities Chinese Health Study. *Environ Int.* oct 2018;119:485-92. doi:10.1016/j.envint.2018.07.016 PubMed PMID: 30048882.
 24. Shanley RP, Hayes RB, Cromar KR, Ito K, Gordon T, Ahn J. Particulate Air Pollution and Clinical Cardiovascular Disease Risk Factors. *Epidemiology.* mars 2016;27(2):291-8. doi:10.1097/EDE.0000000000000426 PubMed PMID: 26605815; PubMed Central PMCID: PMC4959464.
 25. Bell G, Mora S, Greenland P, Tsai M, Gill E, Kaufman JD. Association of Air Pollution Exposures With High-Density Lipoprotein Cholesterol and Particle Number: The Multi-Ethnic Study of Atherosclerosis. *Arterioscler Thromb Vasc Biol.* mai 2017;37(5):976-82. doi:10.1161/ATVBAHA.116.308193 PubMed PMID: 28408373; PubMed Central PMCID: PMC5407952.
 26. Article - Bulletin épidémiologique hebdomadaire [Internet]. [cité 25 sept 2025]. Disponible sur: https://beh.santepubliquefrance.fr/beh/2025/10/2025_10_1.html
 27. Article - Bulletin épidémiologique hebdomadaire [Internet]. [cité 25 sept 2025]. Disponible sur: https://beh.santepubliquefrance.fr/beh/2019/5-6/2019_5-6_2.html
 28. Household air pollution [Internet]. [cité 29 déc 2025]. Disponible sur: <https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health>
 29. Qualité de l'air intérieur | Anses - Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail [Internet]. [cité 28 déc 2025]. Disponible sur: <https://www.anses.fr/fr/content/qualite-de-lair-interieur-1>
 30. Devien L, Giovannelli J, Cuny D, Matran R, Amouyel P, Hulo S, et al. Sources of household air pollution: The association with lung function and respiratory symptoms in middle-aged adult. *Environ Res.* juill 2018;164:140-8. doi:10.1016/j.envres.2018.02.016 PubMed PMID: 29486345.
 31. Curtis L, Lieberman A, Stark M, Rea W, Vetter M. Adverse Health Effects of Indoor Molds. *Journal of Nutritional & Environmental Medicine.* 1 sept 2004;14:261-74. doi:10.1080/13590840400010318
 32. WHO/ILO: Almost 2 million people die from work-related causes each year | International Labour Organization [Internet]. 2021 [cité 30 déc 2025]. Disponible sur: <https://www.ilo.org/resource/news/whoilo-almost-2-million-people-die-work-related-causes-each-year>
 33. SPF. Surveillance épidémiologique des risques professionnels : pourquoi et comment ? Numéro thématique. Risques professionnels : quelle veille sanitaire ? [Internet]. [cité 30 déc 2025]. Disponible sur:

<https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-liees-au-travail/maladies-a-caractere-professionnel/surveillance-epidemiologique-des-risques-professionnels-pourquoi-et-comment-numero-thematique.-risques-professionnels-quelle-veille-sanitaire>

34. Wardyn PM, Edme JL, de Broucker V, Cherot-Kornobis N, Ringeval D, Amouyel P, et al. The impact of occupational exposure to crystalline silica dust on respiratory function (airway obstruction and FEF25-75) in the French general population. *Environ Res.* 1 avr 2023;222:115382. doi:10.1016/j.envres.2023.115382 PubMed PMID: 36736759.
35. Mehta AJ, Miedinger D, Keidel D, Bettschart R, Bircher A, Bridevaux PO, et al. Occupational exposure to dusts, gases, and fumes and incidence of chronic obstructive pulmonary disease in the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults. *Am J Respir Crit Care Med.* 15 juin 2012;185(12):1292-300. doi:10.1164/rccm.201110-1917OC PubMed PMID: 22492989.
36. Lead poisoning [Internet]. [cité 29 déc 2025]. Disponible sur: <https://www.who.int/news-room/fact-sheets/detail/lead-poisoning-and-health>
37. WHO guideline for the clinical management of exposure to lead. Geneva: World Health Organization; 2021. Licence: CC BY-NC-SA 3.0 IGO.
38. WHO guidance to reduce illness due to lead exposure [Internet]. [cité 30 juill 2025]. Disponible sur: <https://www.who.int/news/item/27-10-2021-who-guidance-to-reduce-illness-due-to-lead-exposure>
39. SPF. Imprégnation de la population française par le plomb. Programme national de biosurveillance, Esteban 2014-2016. [Internet]. [cité 21 sept 2025]. Disponible sur: <https://www.santepubliquefrance.fr/import/impregnation-de-la-population-francaise-par-le-plomb.-programme-national-de-biosurveillance-esteban-2014-2016>
40. Jaffe AE, Hyde T, Kleinman J, Weinberg DR, Chenoweth JG, McKay RD, et al. Practical impacts of genomic data « cleaning » on biological discovery using surrogate variable analysis. *BMC Bioinformatics.* 6 nov 2015;16:372. doi:10.1186/s12859-015-0808-5 PubMed PMID: 26545828; PubMed Central PMCID: PMC4636836.
41. Tinkov AA, Filippini T, Ajsuvakova OP, Aaseth J, Gluhcheva YG, Ivanova JM, et al. The role of cadmium in obesity and diabetes. *Science of The Total Environment.* 1 déc 2017;601-602:741-55. doi:10.1016/j.scitotenv.2017.05.224
42. Wani AL, Ara A, Usmani JA. Lead toxicity: a review. *Interdiscip Toxicol.* juin 2015;8(2):55-64. doi:10.1515/intox-2015-0009 PubMed PMID: 27486361; PubMed Central PMCID: PMC4961898.
43. Collin MS, Venkatraman SK, Vijayakumar N, Kanimozhi V, Arbaaz SM, Stacey RGS, et al. Bioaccumulation of lead (Pb) and its effects on human: A review. *Journal of Hazardous Materials Advances.* 1 août 2022;7:100094. doi:10.1016/j.hazadv.2022.100094

44. Havet A, Hulo S, Cuny D, Riant M, Occelli F, Cherot-Kornobis N, et al. Residential exposure to outdoor air pollution and adult lung function, with focus on small airway obstruction. *Environ Res. avr* 2020;183:109161. doi:10.1016/j.envres.2020.109161 PubMed PMID: 32000005.
45. Vineis P, Perera F. Molecular Epidemiology and Biomarkers in Etiologic Cancer Research: The New in Light of the Old. *Cancer Epidemiol Biomarkers Prev.* 11 oct 2007;16(10):1954-65. doi:10.1158/1055-9965.EPI-07-0457
46. Pi-Sunyer X. The medical risks of obesity. *Postgrad Med.* nov 2009;121(6):21-33. doi:10.3810/pgm.2009.11.2074 PubMed PMID: 19940414; PubMed Central PMCID: PMC2879283.
47. de Courrèges A, Occelli F, Muntaner M, Amouyel P, Meirhaeghe A, Dauchet L. The relationship between neighbourhood walkability and cardiovascular risk factors in northern France. *Science of The Total Environment.* 10 juin 2021;772:144877. doi:10.1016/j.scitotenv.2020.144877
48. Riant M, Meirhaeghe A, Giovannelli J, Occelli F, Havet A, Cuny D, et al. Associations between long-term exposure to air pollution, glycosylated hemoglobin, fasting blood glucose and diabetes mellitus in northern France. *Environ Int.* nov 2018;120:121-9. doi:10.1016/j.envint.2018.07.034 PubMed PMID: 30077944.
49. Trouiller-Gerfaux P, Podglajen E, Hulo S, Richeval C, Allorge D, Garat A, et al. The association between blood cadmium and glycated haemoglobin among never-, former, and current smokers: A cross-sectional study in France. *Environ Res.* nov 2019;178:108673. doi:10.1016/j.envres.2019.108673 PubMed PMID: 31520822.
50. Darras-Hostens M, Degrendel M, Amouyel P, Dauchet L. Association between residential air pollution exposure and cardiovascular risk factors in adults living in northern France. *Environ Geochem Health.* 7 juin 2024;46(7):232. doi:10.1007/s10653-024-02006-2 PubMed PMID: 38849665.
51. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* mars 2001;69(3):89-95. doi:10.1067/mcp.2001.113989 PubMed PMID: 11240971.
52. Hubert A, Achour D, Grare C, Zarcone G, Muntaner M, Hamroun A, et al. The relationship between residential exposure to atmospheric pollution and circulating miRNA in adults living in an urban area in northern France. *Environ Int.* avr 2023;174:107913. doi:10.1016/j.envint.2023.107913 PubMed PMID: 37037173.
53. Bentegeac R, Achour D, Grare C, Muntaner M, Gauthier V, Amouyel P, et al. Associations between air pollution and biomarkers of oxidative stress and lung damage in a large population-based sample of non-smoking adults in northern France. *Environ Geochem Health.* 12 avr 2025;47(5):166. doi:10.1007/s10653-025-02472-2 PubMed PMID: 40220195; PubMed Central PMCID: PMC11993482.

54. Darras-Hostens M, Achour D, Muntaner M, Grare C, Zarcone G, Garçon G, et al. Short-term and residential exposure to air pollution: Associations with inflammatory biomarker levels in adults living in northern France. *Sci Total Environ*. 10 août 2022;833:154985. doi:10.1016/j.scitotenv.2022.154985 PubMed PMID: 35398417.
55. Hasin Y, Seldin M, Lusk A. Multi-omics approaches to disease. *Genome Biol*. 5 mai 2017;18(1):83. doi:10.1186/s13059-017-1215-1 PubMed PMID: 28476144; PubMed Central PMCID: PMC5418815.
56. Fiehn O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp Funct Genomics*. 2001;2(3):155-68. doi:10.1002/cfg.82 PubMed PMID: 18628911; PubMed Central PMCID: PMC2447208.
57. Amer B, Deshpande RR, Bird SS. Simultaneous Quantitation and Discovery (SQUAD) Analysis: Combining the Best of Targeted and Untargeted Mass Spectrometry-Based Metabolomics. *Metabolites*. 10 mai 2023;13(5):648. doi:10.3390/metabo13050648 PubMed PMID: 37233689; PubMed Central PMCID: PMC10223305.
58. Miller MG. Environmental Metabolomics: A SWOT Analysis (Strengths, Weaknesses, Opportunities, and Threats). *J Proteome Res*. 1 févr 2007;6(2):540-5. doi:10.1021/pr060623x
59. Fiehn O. Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol*. janv 2002;48(1-2):155-71. PubMed PMID: 11860207.
60. Bonvallet N, Tremblay-Franco M, Chevrier C, Canlet C, Debrauwer L, Cravedi JP, et al. Potential Input From Metabolomics for Exploring and Understanding the Links Between Environment and Health. *Journal of Toxicology and Environmental Health, Part B*. 2 janv 2014;17(1):21-44. doi:10.1080/10937404.2013.860318 PubMed PMID: 24597908.
61. Debik J, Sangermani M, Wang F, Madssen TS, Giskeødegård GF. Multivariate analysis of NMR-based metabolomic data. *NMR Biomed*. févr 2022;35(2):e4638. doi:10.1002/nbm.4638 PubMed PMID: 34738674.
62. Emwas AHM. The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Methods Mol Biol*. 2015;1277:161-93. doi:10.1007/978-1-4939-2377-9_13 PubMed PMID: 25677154.
63. Gloaguen Y, Morton F, Daly R, Gurden R, Rogers S, Wandy J, et al. PiMP my metabolome: an integrated, web-based tool for LC-MS metabolomics data. *Bioinformatics*. 15 déc 2017;33(24):4007-9. doi:10.1093/bioinformatics/btx499 PubMed PMID: 28961954; PubMed Central PMCID: PMC5860087.
64. Rhee J, Loftfield E, Albanes D, Layne TM, Stolzenberg-Solomon R, Liao LM, et al. A metabolomic investigation of serum perfluorooctane sulfonate and perfluorooctanoate. *Environ Int*. oct 2023;180:108198. doi:10.1016/j.envint.2023.108198 PubMed PMID: 37716341; PubMed Central PMCID: PMC10591812.

65. Tessier AJ, Wang F, Liang L, Wittenbecher C, Haslam DE, Eliassen AH, et al. Plasma metabolites of a healthy lifestyle in relation to mortality and longevity: Four prospective US cohort studies. *Med*. 8 mars 2024;5(3):224-238.e5. doi:10.1016/j.medj.2024.01.010 PubMed PMID: 38366602; PubMed Central PMCID: PMC10940196.
66. Alcolea JA, Donat-Vargas C, Chatziioannou AC, Keski-Rahkonen P, Robinot N, Molina AJ, et al. Metabolomic Signatures of Exposure to Nitrate and Trihalomethanes in Drinking Water and Colorectal Cancer Risk in a Spanish Multicentric Study (MCC-Spain). *Environ Sci Technol*. 5 déc 2023;57(48):19316-29. doi:10.1021/acs.est.3c05814 PubMed PMID: 37962559.
67. Theodoridis G, Gika H, Raftery D, Goodacre R, Plumb RS, Wilson ID. Ensuring Fact-Based Metabolite Identification in Liquid Chromatography–Mass Spectrometry-Based Metabolomics. *Anal Chem*. 15 févr 2023;95(8):3909-16. doi:10.1021/acs.analchem.2c05192 PubMed PMID: 36791228; PubMed Central PMCID: PMC9979140.
68. Rousseau K. Développement de nouveaux outils pour l'analyse métabolomique par spectrométrie de masse haute résolution : de l'acquisition de spectres MS/MS pour l'identification large-spectre de métabolites au marquage isotopique pour la quantification [phdthesis] [Internet]. Université Paris-Saclay; 2020 [cité 4 nov 2024]. Disponible sur: <https://theses.hal.science/tel-03793794>
69. Athersuch TJ, Keun HC. Metabolic profiling in human exposome studies. *Mutagenesis*. nov 2015;30(6):755-62. doi:10.1093/mutage/gev060 PubMed PMID: 26290610.
70. Human Metabolome Database [Internet]. [cité 6 nov 2024]. Disponible sur: <https://hmdb.ca/>
71. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, et al. METLIN: a metabolite mass spectral database. *Ther Drug Monit*. déc 2005;27(6):747-51. doi:10.1097/01.ftd.0000179845.53213.39 PubMed PMID: 16404815.
72. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*. sept 2007;3(3):211-21. doi:10.1007/s11306-007-0082-2 PubMed PMID: 24039616; PubMed Central PMCID: PMC3772505.
73. Chemical Entities of Biological Interest (ChEBI) [Internet]. [cité 2 oct 2025]. Disponible sur: <https://www.ebi.ac.uk/chebi/>
74. Wieder C, Frainay C, Poupin N, Rodríguez-Mier P, Vinson F, Cooke J, et al. Pathway analysis in metabolomics: Recommendations for the use of over-representation analysis. *PLoS Comput Biol*. sept 2021;17(9):e1009105. doi:10.1371/journal.pcbi.1009105 PubMed PMID: 34492007; PubMed Central PMCID: PMC8448349.
75. KEGG PATHWAY Database [Internet]. [cité 6 nov 2024]. Disponible sur: <https://www.genome.jp/kegg/pathway.html>

76. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 1 janv 2000;28(1):27-30. doi:10.1093/nar/28.1.27 PubMed PMID: 10592173; PubMed Central PMCID: PMC102409.
77. Lehninger. *Lehninger Principles of Biochemistry*. chapter 1.
78. Biological Pathways Fact Sheet [Internet]. [cité 12 déc 2024]. Disponible sur: <https://www.genome.gov/about-genomics/fact-sheets/Biological-Pathways-Fact-Sheet>
79. Xia J, Sinelnikov IV, Han B, Wishart DS. MetaboAnalyst 3.0--making metabolomics more meaningful. *Nucleic Acids Res.* 1 juill 2015;43(W1):W251-257. doi:10.1093/nar/gkv380 PubMed PMID: 25897128; PubMed Central PMCID: PMC4489235.
80. Wishart DS. Advances in metabolite identification. *Bioanalysis.* août 2011;3(15):1769-82. doi:10.4155/bio.11.155 PubMed PMID: 21827274.
81. Lü L, Pan L, Zhou T, Zhang YC, Stanley HE. Toward link predictability of complex networks. *Proc Natl Acad Sci U S A.* 24 févr 2015;112(8):2325-30. doi:10.1073/pnas.1424644112 PubMed PMID: 25659742; PubMed Central PMCID: PMC4345601.
82. Robinson JL, Kocabaş P, Wang H, Cholley PE, Cook D, Nilsson A, et al. An atlas of human metabolism. *Sci Signal.* 24 mars 2020;13(624):eaaz1482. doi:10.1126/scisignal.aaz1482 PubMed PMID: 32209698; PubMed Central PMCID: PMC7331181.
83. Frainay C, Jourdan F. Computational methods to identify metabolic sub-networks based on metabolomic profiles. *Brief Bioinform.* janv 2017;18(1):43-56. doi:10.1093/bib/bbv115 PubMed PMID: 26822099.
84. Misra BB, Langefeld C, Olivier M, Cox LA. Integrated omics: tools, advances and future approaches [Internet]. 1 janv 2019. doi:10.1530/JME-18-0055
85. Yamada R, Okada D, Wang J, Basak T, Koyama S. Interpretation of omics data analyses. *J Hum Genet.* janv 2021;66(1):93-102. doi:10.1038/s10038-020-0763-5 PubMed PMID: 32385339; PubMed Central PMCID: PMC7728595.
86. Clish CB. Metabolomics: an emerging but powerful tool for precision medicine. *Cold Spring Harb Mol Case Stud.* oct 2015;1(1):a000588. doi:10.1101/mcs.a000588 PubMed PMID: 27148576; PubMed Central PMCID: PMC4850886.
87. Chen Y, Li EM, Xu LY. Guide to Metabolomics Analysis: A Bioinformatics Workflow. *Metabolites.* 15 avr 2022;12(4):357. doi:10.3390/metabo12040357 PubMed PMID: 35448542; PubMed Central PMCID: PMC9032224.
88. Feng B, Liu C, Yi T, Song X, Wang Y, Liu S, et al. Perturbation of amino acid metabolism mediates air pollution associated vascular dysfunction in healthy adults. *Environ Res.* oct 2021;201:111512. doi:10.1016/j.envres.2021.111512 PubMed PMID: 34166659.

89. Li GX, Duan YY, Wang Y, Bian LJ, Xiong MR, Song WP, et al. Potential urinary biomarkers in young adults with short-term exposure to particulate matter and bioaerosols identified using an unbiased metabolomic approach. *Environ Pollut*. 15 juill 2022;305:119308. doi:10.1016/j.envpol.2022.119308 PubMed PMID: 35443204.
90. Cruz R, Pasqua L, Silveira A, Damasceno M, Matsuda M, Martins M, et al. Traffic-related air pollution and endurance exercise: Characterizing non-targeted serum metabolomics profiling. *Environ Pollut*. 15 déc 2021;291:118204. doi:10.1016/j.envpol.2021.118204 PubMed PMID: 34560574.
91. Quach A, Giovannelli J, Chérot-Kornobis N, Ciuchete A, Clément G, Matran R, et al. Prevalence and underdiagnosis of airway obstruction among middle-aged adults in northern France: The ELISABET study 2011-2013. *Respir Med*. déc 2015;109(12):1553-61. doi:10.1016/j.rmed.2015.10.012 PubMed PMID: 26564001.
92. ATMO Nord-Pas-de-Calais. Bilan territorial 2015 - Communauté Urbaine de Dunkerque [Internet]. 2016 [cited 2018 Jul 11]. Available from: http://www.atmo-hdf.fr/joomlatools-files/docman-files/BT2015_CUD_VF.compressed.pdf.
93. ATMO Nord-Pas-de-Calais. Bilan territorial 2015 - Métropole Européenne de Lille [Internet]. 2016 [cited 2018 Jul 11]. Available from: http://www.atmo-hdf.fr/joomlatools-files/docman-files/Bilans-territoriaux/BT2015_MEL_VF.pdf.
94. RAPENEAU P, PATRIS J. COMMUNAUTÉ URBAINE D'ARRAS. 2014;8.
95. Thévenot EA, Roux A, Xu Y, Ezan E, Junot C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J Proteome Res*. 7 août 2015;14(8):3322-35. doi:10.1021/acs.jproteome.5b00354 PubMed PMID: 26088811.
96. Thévenot EA, Roux A, Xu Y, Ezan E, Junot C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J Proteome Res*. 7 août 2015;14(8):3322-35. doi:10.1021/acs.jproteome.5b00354 PubMed PMID: 26088811.
97. Brook RD, Rajagopalan S, Pope CA, Brook JR, Bhatnagar A, Diez-Roux AV, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*. 1 juin 2010;121(21):2331-78. doi:10.1161/CIR.0b013e3181d8e111 PubMed PMID: 20458016.
98. Li R, Navab M, Pakbin P, Ning Z, Navab K, Hough G, et al. Ambient ultrafine particles alter lipid metabolism and HDL anti-oxidant capacity in LDLR-null mice. *Journal of Lipid Research*. juin 2013;54(6):6. doi:10.1194/jlr.M035014
99. Rao X, Zhong J, Maiseyeu A, Gopalakrishnan B, Villamena FA, Chen LC, et al. CD36-Dependent 7-Ketocholesterol Accumulation in Macrophages Mediates Progression of Atherosclerosis in Response to Chronic Air Pollution Exposure Novelty and Significance. *Circulation Research*. 10 oct 2014;115(9):9. doi:10.1161/CIRCRESAHA.115.304666 PubMed PMID: 25186795.

100. Bai Y, Sun Q. Fine particulate matter air pollution and atherosclerosis: Mechanistic insights. *Biochimica et Biophysica Acta (BBA) - General Subjects*. déc 2016;1860(12):12. doi:10.1016/j.bbagen.2016.04.030
101. Gaio V, Roquette R, Dias CM, Nunes B. Ambient air pollution and lipid profile: Systematic review and meta-analysis. *Environ Pollut*. nov 2019;254(Pt B):113036. doi:10.1016/j.envpol.2019.113036 PubMed PMID: 31465899.
102. Cai Y, Hansell AL, Blangiardo M, Burton PR, BioSHaRE, de Hoogh K, et al. Long-term exposure to road traffic noise, ambient air pollution, and cardiovascular risk factors in the HUNT and lifelines cohorts. *Eur Heart J*. 1 août 2017;38(29):2290-6. doi:10.1093/eurheartj/ehx263 PubMed PMID: 28575405; PubMed Central PMCID: PMC5837618.
103. Shin J, Choi J, Kim KJ. Association between long-term exposure of ambient air pollutants and cardiometabolic diseases: A 2012 Korean Community Health Survey. *Nutr Metab Cardiovasc Dis*. févr 2019;29(2):144-51. doi:10.1016/j.numecd.2018.09.008 PubMed PMID: 30595346.
104. McGuinn LA, Schneider A, McGarrah RW, Ward-Caviness C, Neas LM, Di Q, et al. Association of long-term PM2.5 exposure with traditional and novel lipid measures related to cardiovascular disease risk. *Environ Int*. janv 2019;122:193-200. doi:10.1016/j.envint.2018.11.001 PubMed PMID: 30446244; PubMed Central PMCID: PMC6467069.
105. Kim KN, Ha B, Seog W, Hwang IU. Long-term exposure to air pollution and the blood lipid levels of healthy young men. *Environ Int*. mars 2022;161:107119. doi:10.1016/j.envint.2022.107119 PubMed PMID: 35123376.
106. He ZZ, Guo PY, Xu SL, Zhou Y, Jalaludin B, Leskinen A, et al. Associations of Particulate Matter Sizes and Chemical Constituents with Blood Lipids: A Panel Study in Guangzhou, China. *Environ Sci Technol*. 20 avr 2021;55(8):5065-75. doi:10.1021/acs.est.0c06974 PubMed PMID: 33764049.
107. Gui ZH, Yang BY, Zou ZY, Ma J, Jing J, Wang HJ, et al. Exposure to ambient air pollution and blood lipids in children and adolescents: A national population based study in China. *Environ Pollut*. nov 2020;266(Pt 3):115422. doi:10.1016/j.envpol.2020.115422 PubMed PMID: 32829032.
108. Yang BY, Qian Z, Howard SW, Vaughn MG, Fan SJ, Liu KK, et al. Global association between ambient air pollution and blood pressure: A systematic review and meta-analysis. *Environ Pollut*. avr 2018;235:576-88. doi:10.1016/j.envpol.2018.01.001 PubMed PMID: 29331891.
109. Lin H, Guo Y, Zheng Y, Di Q, Liu T, Xiao J, et al. Long-Term Effects of Ambient PM2.5 on Hypertension and Blood Pressure and Attributable Risk Among Older Chinese Adults. *Hypertension*. mai 2017;69(5):806-12. doi:10.1161/HYPERTENSIONAHA.116.08839 PubMed PMID: 28348017; PubMed Central PMCID: PMC5447129.

110. Liu C, Chen R, Zhao Y, Ma Z, Bi J, Liu Y, et al. Associations between ambient fine particulate air pollution and hypertension: A nationwide cross-sectional study in China. *Sci Total Environ.* 15 avr 2017;584-585:869-74. doi:10.1016/j.scitotenv.2017.01.133 PubMed PMID: 28153400; PubMed Central PMCID: PMC8102884.
111. Bilenko N, van Rossem L, Brunekreef B, Beelen R, Eeftens M, Hoek G, et al. Traffic-related air pollution and noise and children's blood pressure: results from the PIAMA birth cohort study. *Eur J Prev Cardiol.* janv 2015;22(1):4-12. doi:10.1177/2047487313505821 PubMed PMID: 24047569.
112. Harrabi I, Rondeau V, Dartigues JF, Tessier JF, Filleul L. Effects of particulate air pollution on systolic blood pressure: A population-based approach. *Environ Res.* mai 2006;101(1):89-93. doi:10.1016/j.envres.2006.01.012 PubMed PMID: 16563371.
113. Zhang J, Zhang F, Xin C, Duan Z, Wei J, Zhang X, et al. Associations of long-term exposure to air pollution, physical activity with blood pressure and prevalence of hypertension: the China Health and Retirement Longitudinal Study. *Front Public Health.* 3 mai 2023;11:1137118. doi:10.3389/fpubh.2023.1137118 PubMed PMID: 37206865; PubMed Central PMCID: PMC10189054.
114. *Annales du Contrôle National de Qualité des Analyses de Biologie Médicale* [Internet]. [cité 30 août 2018]. Disponible sur: https://ansm.sante.fr/var/ansm_site/storage/original/application/cf2f308ef14ebd9ca18fc7041d844655.pdf
115. Miller KA, Siscovick DS, Sheppard L, Shepherd K, Sullivan JH, Anderson GL, et al. Long-Term Exposure to Air Pollution and Incidence of Cardiovascular Events in Women. *New England Journal of Medicine.* 1 févr 2007;356(5):5. doi:10.1056/NEJMoa054409 PubMed PMID: 17267905.
116. Clement G, Giovannelli J, Cottel D, Montaye M, Ciuchete A, Dallongeville J, et al. Changes over time in the prevalence and treatment of cardiovascular risk factors, and contributions to time trends in coronary mortality over 25 years in the Lille urban area (northern France). *Arch Cardiovasc Dis.* déc 2017;110(12):689-99. doi:10.1016/j.acvd.2017.03.009 PubMed PMID: 28958407.
117. Cui JS, Hopper JL, Harrap SB. Antihypertensive treatments obscure familial contributions to blood pressure variation. *Hypertension.* févr 2003;41(2):207-10. doi:10.1161/01.hyp.0000044938.94050.e3 PubMed PMID: 12574083.
118. Saber AT, Jacobsen NR, Jackson P, Poulsen SS, Kyjovska ZO, Halappanavar S, et al. Particle-induced pulmonary acute phase response may be the causal link between particle inhalation and cardiovascular disease: Particle inhalation and cardiovascular disease. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology.* nov 2014;6(6):6. doi:10.1002/wnan.1279
119. Bourdon JA, Halappanavar S, Saber AT, Jacobsen NR, Williams A, Wallin H, et al. Hepatic and Pulmonary Toxicogenomic Profiles in Mice Intratracheally Instilled With Carbon Black Nanoparticles Reveal Pulmonary Inflammation, Acute Phase Response, and

- Alterations in Lipid Homeostasis. *Toxicol Sci.* juin 2012;127(2):2. doi:10.1093/toxsci/kfs119 PubMed PMID: 22461453; PubMed Central PMCID: PMC3355316.
120. Tsutsumi K. Lipoprotein lipase and atherosclerosis. *Curr Vasc Pharmacol.* mars 2003;1(1):11-7. doi:10.2174/1570161033386673 PubMed PMID: 15320848.
121. Cai Y, Zhang B, Ke W, Feng B, Lin H, Xiao J, et al. Associations of Short-Term and Long-Term Exposure to Ambient Air Pollutants With Hypertension: A Systematic Review and Meta-Analysis. *Hypertension.* juill 2016;68(1):62-70. doi:10.1161/HYPERTENSIONAHA.116.07218 PubMed PMID: 27245182.
122. Perez CM, Hazari MS, Farraj AK. Role of autonomic reflex arcs in cardiovascular responses to air pollution exposure. *Cardiovasc Toxicol.* janv 2015;15(1):69-78. doi:10.1007/s12012-014-9272-0 PubMed PMID: 25123706; PubMed Central PMCID: PMC4766835.
123. Bellavia A, Urch B, Speck M, Brook RD, Scott JA, Albeti B, et al. DNA hypomethylation, ambient particulate matter, and increased blood pressure: findings from controlled human exposure experiments. *J Am Heart Assoc.* 19 juin 2013;2(3):e000212. doi:10.1161/JAHA.113.000212 PubMed PMID: 23782920; PubMed Central PMCID: PMC3698788.
124. Rothman KJ, Greenland S, Lash TL. Chapter 8, Precision and Validity in Epidemiologic Studies, *Modern Epidemiology.* Lippincott Williams & Wilkins; 2008. 776 p.
125. Air pollution [Internet]. [cité 19 mai 2021]. Disponible sur: <https://www.who.int/news-room/air-pollution>
126. Pearson TA, Mensah GA, Alexander RW, Anderson JL, Cannon RO, Criqui M, et al. Markers of inflammation and cardiovascular disease: application to clinical and public health practice: A statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. *Circulation.* 28 janv 2003;107(3):499-511. doi:10.1161/01.cir.0000052939.59093.45 PubMed PMID: 12551878.
127. León-Pedroza JI, González-Tapia LA, del Olmo-Gil E, Castellanos-Rodríguez D, Escobedo G, González-Chávez A. Inflamación sistémica de grado bajo y su relación con el desarrollo de enfermedades metabólicas: de la evidencia molecular a la aplicación clínica. *Cirugía y Cirujanos.* 1 nov 2015;83(6):543-51. doi:10.1016/j.circir.2015.05.041
128. Lee HM, Cha JM, Lee JL, Jeon JW, Shin HP, Joo KR, et al. High C-reactive protein level is associated with high-risk adenoma. *Intest Res.* oct 2017;15(4):511-7. doi:10.5217/ir.2017.15.4.511 PubMed PMID: 29142519; PubMed Central PMCID: PMC5683982.
129. Leonard BE. Inflammation and depression: a causal or coincidental link to the pathophysiology? *Acta Neuropsychiatr.* févr 2018;30(1):1-16. doi:10.1017/neu.2016.69 PubMed PMID: 28112061.

130. Green R, Broadwin R, Malig B, Basu R, Gold EB, Qi L, et al. Long- and Short-term Exposure to Air Pollution and Inflammatory/Hemostatic Markers in Midlife Women. *Epidemiology*. mars 2016;27(2):211-20. doi:10.1097/EDE.0000000000000421 PubMed PMID: 26600256; PubMed Central PMCID: PMC4841679.
131. Rioux CL, Tucker KL, Mwamburi M, Gute DM, Cohen SA, Brugge D. Residential traffic exposure, pulse pressure, and C-reactive protein: consistency and contrast among exposure characterization methods. *Environ Health Perspect*. juin 2010;118(6):803-11. doi:10.1289/ehp.0901182 PubMed PMID: 20123638; PubMed Central PMCID: PMC2898857.
132. Rifai N, Ridker PM. High-sensitivity C-reactive protein: a novel and promising marker of coronary heart disease. *Clin Chem*. mars 2001;47(3):403-11. PubMed PMID: 11238289.
133. Barnes PJ. Cellular and molecular mechanisms of asthma and COPD. *Clin Sci (Lond)*. 1 juill 2017;131(13):1541-58. doi:10.1042/CS20160487 PubMed PMID: 28659395.
134. Akiki Z, Rava M, Diaz Gil O, Pin I, le Moual N, Siroux V, et al. Serum cytokine profiles as predictors of asthma control in adults from the EGEA study. *Respir Med*. avr 2017;125:57-64. doi:10.1016/j.rmed.2017.03.002 PubMed PMID: 28340863.
135. Shamsollahi HR, Jahanbin B, Rafieian S, Yunesian M. Particulates induced lung inflammation and its consequences in the development of restrictive and obstructive lung diseases: a systematic review. *Environ Sci Pollut Res Int*. mai 2021;28(20):25035-50. doi:10.1007/s11356-021-13559-5 PubMed PMID: 33779901.
136. Plé C, Fan Y, Ait Yahia S, Vorng H, Everaere L, Chenivesse C, et al. Polycyclic Aromatic Hydrocarbons Reciprocally Regulate IL-22 and IL-17 Cytokines in Peripheral Blood Mononuclear Cells from Both Healthy and Asthmatic Subjects. *PLoS One*. 10 avr 2015;10(4):e0122372. doi:10.1371/journal.pone.0122372 PubMed PMID: 25860963; PubMed Central PMCID: PMC4393221.
137. Pope CA, Hansen ML, Long RW, Nielsen KR, Eatough NL, Wilson WE, et al. Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environ Health Perspect*. mars 2004;112(3):339-45. doi:10.1289/ehp.6588 PubMed PMID: 14998750; PubMed Central PMCID: PMC1241864.
138. Westberg H, Elihn K, Andersson E, Persson B, Andersson L, Bryngelsson IL, et al. Inflammatory markers and exposure to airborne particles among workers in a Swedish pulp and paper mill. *Int Arch Occup Environ Health*. juill 2016;89(5):813-22. doi:10.1007/s00420-016-1119-5 PubMed PMID: 26875192; PubMed Central PMCID: PMC4871919.
139. Shakya KM, Peltier RE, Zhang Y, Pandey BD. Roadside Exposure and Inflammation Biomarkers among a Cohort of Traffic Police in Kathmandu, Nepal. *Int J Environ Res Public Health*. 29 janv 2019;16(3). doi:10.3390/ijerph16030377 PubMed PMID: 30699969; PubMed Central PMCID: PMC6388290.

140. Riaz H, Syed BM, Laghari Z, Pirzada S. Analysis of inflammatory markers in apparently healthy automobile vehicle drivers in response to exposure to traffic pollution fumes. *Pak J Med Sci.* juin 2020;36(4):657-62. doi:10.12669/pjms.36.4.2025 PubMed PMID: 32494251; PubMed Central PMCID: PMC7260889.
141. Huang WH, Yen TH, Chan MJ, Su YJ. Environmental carbon monoxide level is associated with the level of high-sensitivity C-reactive protein in peritoneal dialysis patients. *Medicine (Baltimore).* nov 2014;93(26):e181. doi:10.1097/MD.000000000000181 PubMed PMID: 25474434; PubMed Central PMCID: PMC4616395.
142. Sajid Jabbar A, Ali ET. Impact of Petroleum Exposure on Some Hematological Indices, Interleukin-6, and Inflammatory Markers of Workers at Petroleum Stations in Basra City. *J Environ Public Health.* 2020;2020:7693891. doi:10.1155/2020/7693891 PubMed PMID: 32831856; PubMed Central PMCID: PMC7424534.
143. Brucker N, Moro AM, Charão MF, Durgante J, Freitas F, Baierle M, et al. Biomarkers of occupational exposure to air pollution, inflammation and oxidative damage in taxi drivers. *Sci Total Environ.* 1 oct 2013;463-464:884-93. doi:10.1016/j.scitotenv.2013.06.098 PubMed PMID: 23872245.
144. Dai Y, Zhang X, Zhang R, Zhao X, Duan H, Niu Y, et al. Long-term exposure to diesel engine exhaust affects cytokine expression among occupational population. *Toxicol Res (Camb).* 1 mars 2016;5(2):674-81. doi:10.1039/c5tx00462d PubMed PMID: 30090380; PubMed Central PMCID: PMC6060680.
145. Bassig BA, Dai Y, Vermeulen R, Ren D, Hu W, Duan H, et al. Occupational exposure to diesel engine exhaust and alterations in immune/inflammatory markers: a cross-sectional molecular epidemiology study in China. *Carcinogenesis.* 26 oct 2017;38(11):1104-11. doi:10.1093/carcin/bgx081 PubMed PMID: 28968774; PubMed Central PMCID: PMC5862277.
146. Negherbon JP, Romero K, Williams DL, Guerrero-Preston RE, Hartung T, Scott AL, et al. Whole Blood Cytokine Response to Local Traffic-Related Particulate Matter in Peruvian Children With and Without Asthma. *Front Pharmacol.* 2017;8:157. doi:10.3389/fphar.2017.00157 PubMed PMID: 28424616; PubMed Central PMCID: PMC5371665.
147. Buxton MA, Meraz-Cruz N, Sanchez BN, Gronlund CJ, Foxman B, Vadillo-Ortega F, et al. Air pollution and inflammation: Findings from concurrent repeated measures of systemic and reproductive tract cytokines during term pregnancy in Mexico City. *Sci Total Environ.* 1 sept 2019;681:235-41. doi:10.1016/j.scitotenv.2019.05.041 PubMed PMID: 31103661; PubMed Central PMCID: PMC6582973.
148. Chen R, Li H, Cai J, Wang C, Lin Z, Liu C, et al. Fine Particulate Air Pollution and the Expression of microRNAs and Circulating Cytokines Relevant to Inflammation, Coagulation, and Vasoconstriction. *Environ Health Perspect.* 17 janv 2018;126(1):017007. doi:10.1289/EHP1447 PubMed PMID: 29342453; PubMed Central PMCID: PMC6014692.

149. Klümper C, Krämer U, Lehmann I, von Berg A, Berdel D, Herberth G, et al. Air pollution and cytokine responsiveness in asthmatic and non-asthmatic children. *Environ Res. avr* 2015;138:381-90. doi:10.1016/j.envres.2015.02.034 PubMed PMID: 25769127.
150. Taj T, Gliga AR, Hedmer M, Wahlberg K, Assarsson E, Lundh T, et al. Effect of welding fumes on the cardiovascular system: a six-year longitudinal study. *Scand J Work Environ Health. 1 janv 2021*;47(1):52-61. doi:10.5271/sjweh.3908 PubMed PMID: 32725248; PubMed Central PMCID: PMC7801138.
151. Baumann R, Joraslafsky S, Markert A, Rack I, Davatgarbenam S, Kossack V, et al. IL-6, a central acute-phase mediator, as an early biomarker for exposure to zinc-based metal fumes. *Toxicology. 12 déc 2016*;373:63-73. doi:10.1016/j.tox.2016.11.001 PubMed PMID: 27816692.
152. Pilz V, Wolf K, Breitner S, Ruckerl R, Koenig W, Rathmann W, et al. C-reactive protein (CRP) and long-term air pollution with a focus on ultrafine particles. *Int J Hyg Environ Health. avr 2018*;221(3):510-8. doi:10.1016/j.ijheh.2018.01.016 PubMed PMID: 29428699.
153. Forbes LJJ, Patel MD, Rudnicka AR, Cook DG, Bush T, Stedman JR, et al. Chronic exposure to outdoor air pollution and markers of systemic inflammation. *Epidemiology. mars 2009*;20(2):245-53. doi:10.1097/EDE.0b013e318190ea3f PubMed PMID: 19234416.
154. Matsuda M, Bonatti R, Marquezini MV, Garcia MLB, Santos UP, Braga ALF, et al. Lacrimal Cytokines Assessment in Subjects Exposed to Different Levels of Ambient Air Pollution in a Large Metropolitan Area. *PLoS One. 2015*;10(11):e0143131. doi:10.1371/journal.pone.0143131 PubMed PMID: 26588473; PubMed Central PMCID: PMC4654582.
155. Dai Y, Ren D, Bassig BA, Vermeulen R, Hu W, Niu Y, et al. Occupational exposure to diesel engine exhaust and serum cytokine levels. *Environ Mol Mutagen. mars 2018*;59(2):144-50. doi:10.1002/em.22142 PubMed PMID: 29023999; PubMed Central PMCID: PMC6438621.
156. Young BN, Peel JL, Nelson TL, Bachand AM, Heiderscheidt JM, Luna B, et al. C-reactive protein from dried blood spots: Application to household air pollution field studies. *Indoor Air. janv 2020*;30(1):24-30. doi:10.1111/ina.12603 PubMed PMID: 31539172; PubMed Central PMCID: PMC6917940.
157. van den Hooven EH, de Kluizenaar Y, Pierik FH, Hofman A, van Ratingen SW, Zandveld PYJ, et al. Chronic air pollution exposure during pregnancy and maternal and fetal C-reactive protein levels: the Generation R Study. *Environ Health Perspect. mai 2012*;120(5):746-51. doi:10.1289/ehp.1104345 PubMed PMID: 22306530; PubMed Central PMCID: PMC3346784.
158. Latzin P, Frey U, Armann J, Kieninger E, Fuchs O, Rössli M, et al. Exposure to moderate air pollution during late pregnancy and cord blood cytokine secretion in healthy neonates. *PLoS One. 2011*;6(8):e23130. doi:10.1371/journal.pone.0023130 PubMed PMID: 21826232; PubMed Central PMCID: PMC3149643.

159. Gruzieva O, Merid SK, Gref A, Gajulapuri A, Lemonnier N, Ballereau S, et al. Exposure to Traffic-Related Air Pollution and Serum Inflammatory Cytokines in Children. *Environ Health Perspect*. 16 juin 2017;125(6):067007. doi:10.1289/EHP460 PubMed PMID: 28669936; PubMed Central PMCID: PMC5714301.
160. Elbarbary M, Oganesyanyan A, Honda T, Morgan G, Guo Y, Guo Y, et al. Systemic Inflammation (C-Reactive Protein) in Older Chinese Adults Is Associated with Long-Term Exposure to Ambient Air Pollution. *Int J Environ Res Public Health*. 22 mars 2021;18(6):3258. doi:10.3390/ijerph18063258 PubMed PMID: 33809857; PubMed Central PMCID: PMC8004276.
161. Lane KJ, Levy JI, Scammell MK, Peters JL, Patton AP, Reisner E, et al. Association of modeled long-term personal exposure to ultrafine particles with inflammatory and coagulation biomarkers. *Environ Int*. août 2016;92-93:173-82. doi:10.1016/j.envint.2016.03.013 PubMed PMID: 27107222; PubMed Central PMCID: PMC4902720.
162. Tsai DH, Riediker M, Berchet A, Paccaud F, Waeber G, Vollenweider P, et al. Effects of short- and long-term exposures to particulate matter on inflammatory marker levels in the general population. *Environ Sci Pollut Res Int*. juill 2019;26(19):19697-704. doi:10.1007/s11356-019-05194-y PubMed PMID: 31079306.
163. Hennig F, Fuks K, Moebus S, Weinmayr G, Memmesheimer M, Jakobs H, et al. Association between Source-Specific Particulate Matter Air Pollution and hs-CRP: Local Traffic and Industrial Emissions. *Environ Health Perspect*. juill 2014;122(7):703-10. doi:10.1289/ehp.1307081 PubMed PMID: 24755038; PubMed Central PMCID: PMC4080540.
164. Pascal M, Corso M, Ung A, Declercq C, Medina S. Guidelines for assessing the health impacts of air pollution in European cities Work Package 5 Deliverable D5. 40.
165. CERC > Environmental software > ADMS-Urban model [Internet]. [cité 28 déc 2021]. Disponible sur: <http://www.cerc.co.uk/environmental-software/ADMS-Urban-model.html>
166. NUMTECH - Logiciels gamme ADMS - Conseils, vente exclusive, formation et assistance [Internet]. [cité 28 déc 2021]. Disponible sur: <http://www.numtech.fr/log.php?rub=urban>
167. Babak O, Deutsch CV. Statistical approach to inverse distance interpolation. *Stoch Environ Res Risk Assess*. 1 juill 2009;23(5):543-53. doi:10.1007/s00477-008-0226-6
168. Xiong Y, Du K. Source-resolved attribution of ground-level ozone formation potential from VOC emissions in Metropolitan Vancouver, BC. *Science of The Total Environment*. 15 juin 2020;721:137698. doi:10.1016/j.scitotenv.2020.137698
169. van Eeden SF, Tan WC, Suwa T, Mukae H, Terashima T, Fujii T, et al. Cytokines involved in the systemic inflammatory response induced by exposure to particulate matter air pollutants (PM(10)). *Am J Respir Crit Care Med*. 1 sept 2001;164(5):826-30. doi:10.1164/ajrccm.164.5.2010160 PubMed PMID: 11549540.

170. Becker S, Mundandhara S, Devlin RB, Madden M. Regulation of cytokine production in human alveolar macrophages and airway epithelial cells in response to ambient air pollution particles: further mechanistic studies. *Toxicol Appl Pharmacol.* 1 sept 2005;207(2 Suppl):269-75. doi:10.1016/j.taap.2005.01.023 PubMed PMID: 15993911.
171. Glencross DA, Ho TR, Camiña N, Hawrylowicz CM, Pfeffer PE. Air pollution and its effects on the immune system. *Free Radical Biology and Medicine.* mai 2020;151:56-68. doi:10.1016/j.freeradbiomed.2020.01.179
172. Gabay C, Kushner I. Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med.* 11 févr 1999;340(6):448-54. doi:10.1056/NEJM199902113400607 PubMed PMID: 9971870.
173. Szalai AJ, Agrawal A, Greenhough TJ, Volanakis JE. C-reactive protein: structural biology and host defense function. *Clin Chem Lab Med.* mars 1999;37(3):265-70. doi:10.1515/CCLM.1999.046 PubMed PMID: 10353470.
174. Sheriff A, Kayser S, Brunner P, Vogt B. C-Reactive Protein Triggers Cell Death in Ischemic Cells. *Front Immunol.* 10 févr 2021;12:630430. doi:10.3389/fimmu.2021.630430 PubMed PMID: 33679775; PubMed Central PMCID: PMC7934421.
175. Watterson TL, Hamilton B, Martin R, Coulombe RA. Urban particulate matter causes ER stress and the unfolded protein response in human lung cells. *Toxicol Sci.* nov 2009;112(1):111-22. doi:10.1093/toxsci/kfp186 PubMed PMID: 19675143.
176. Ramage L, Guy K. Expression of C-reactive protein and heat-shock protein-70 in the lung epithelial cell line A549, in response to PM10 exposure. *Inhal Toxicol.* juin 2004;16(6-7):447-52. doi:10.1080/08958370490439614 PubMed PMID: 15204760.
177. Ramage L, Proudfoot L, Guy K. Expression of C-reactive protein in human lung epithelial cells and upregulation by cytokines and carbon particles. *Inhal Toxicol.* août 2004;16(9):607-13. doi:10.1080/08958370490464599 PubMed PMID: 16036753.
178. Vogel CFA, Sciuillo E, Wong P, Kuzmicky P, Kado N, Matsumura F. Induction of proinflammatory cytokines and C-reactive protein in human macrophage cell line U937 exposed to air pollution particulates. *Environ Health Perspect.* nov 2005;113(11):1536-41. doi:10.1289/ehp.8094 PubMed PMID: 16263508; PubMed Central PMCID: PMC1310915.
179. Niwa Y, Hiura Y, Sawamura H, Iwai N. Inhalation exposure to carbon black induces inflammatory response in rats. *Circ J.* janv 2008;72(1):144-9. doi:10.1253/circj.72.144 PubMed PMID: 18159116.
180. Rohr AC, Wagner JG, Morishita M, Kamal A, Keeler GJ, Harkema JR. Cardiopulmonary responses in spontaneously hypertensive and Wistar-Kyoto rats exposed to concentrated ambient particles from Detroit, Michigan. *Inhal Toxicol.* mai 2010;22(6):522-33. doi:10.3109/08958370903524509 PubMed PMID: 20384466.
181. Upadhyay S, Ganguly K, Stoeger T, Semmler-Bhenke M, Takenaka S, Kreyling WG, et al. Cardiovascular and inflammatory effects of intratracheally instilled ambient dust from

- Augsburg, Germany, in spontaneously hypertensive rats (SHRs). *Part Fibre Toxicol.* 29 sept 2010;7:27. doi:10.1186/1743-8977-7-27 PubMed PMID: 20920269; PubMed Central PMCID: PMC2956709.
182. Lei YC, Hwang JS, Chan CC, Lee CT, Cheng TJ. Enhanced oxidative stress and endothelial dysfunction in streptozotocin-diabetic rats exposed to fine particles. *Environ Res.* nov 2005;99(3):335-43. doi:10.1016/j.envres.2005.03.011 PubMed PMID: 16307975.
183. Heijink IH, Brandenburg SM, Postma DS, van Oosterhout AJM. Cigarette smoke impairs airway epithelial barrier function and cell-cell contact recovery. *Eur Respir J.* févr 2012;39(2):419-28. doi:10.1183/09031936.00193810 PubMed PMID: 21778164.
184. Jones JG, Minty BD, Lawler P, Hulands G, Crawley JC, Veall N. Increased alveolar epithelial permeability in cigarette smokers. *Lancet.* 12 janv 1980;1(8159):66-8. doi:10.1016/s0140-6736(80)90493-6 PubMed PMID: 6101416.
185. Danov O, Wolff M, Bartel S, Böhlen S, Obernolte H, Wronski S, et al. Cigarette Smoke Affects Dendritic Cell Populations, Epithelial Barrier Function, and the Immune Response to Viral Infection With H1N1. *Front Med (Lausanne).* 2020;7:571003. doi:10.3389/fmed.2020.571003 PubMed PMID: 33240904; PubMed Central PMCID: PMC7678748.
186. Mehta H, Nazzal K, Sadikot RT. Cigarette smoking and innate immunity. *Inflamm Res.* nov 2008;57(11):497-503. doi:10.1007/s00011-008-8078-6 PubMed PMID: 19109742.
187. Dietrich M, Block G, Norkus EP, Hudes M, Traber MG, Cross CE, et al. Smoking and exposure to environmental tobacco smoke decrease some plasma antioxidants and increase gamma-tocopherol in vivo after adjustment for dietary antioxidant intakes. *Am J Clin Nutr.* janv 2003;77(1):160-6. doi:10.1093/ajcn/77.1.160 PubMed PMID: 12499336.
188. Hemilä H. The effect of β -carotene on the mortality of male smokers is modified by smoking and by vitamins C and E: evidence against a uniform effect of nutrient. *J Nutr Sci.* 11 mars 2020;9:e11. doi:10.1017/jns.2020.3 PubMed PMID: 32215208; PubMed Central PMCID: PMC7082716.
189. Münzel T, Sørensen M, Gori T, Schmidt FP, Rao X, Brook FR, et al. Environmental stressors and cardio-metabolic disease: part II-mechanistic insights. *Eur Heart J.* 21 févr 2017;38(8):557-64. doi:10.1093/eurheartj/ehw294 PubMed PMID: 27460891; PubMed Central PMCID: PMC5381593.
190. Sun S, Cao W, Chan KP, Ran J, Ge Y, Zhang Y, et al. Cigarette smoking increases deaths associated with air pollution in Hong Kong. *Atmospheric Environment.* 15 févr 2020;223:117266. doi:10.1016/j.atmosenv.2020.117266
191. Panasevich S, Leander K, Rosenlund M, Ljungman P, Bellander T, de Faire U, et al. Associations of long- and short-term air pollution exposure with markers of inflammation and coagulation in a population sample. *Occup Environ Med.* nov 2009;66(11):747-53. doi:10.1136/oem.2008.043471 PubMed PMID: 19687019.

192. Gentile M, Panico S, Rubba F, Mattiello A, Chiodini P, Jossa F, et al. Obesity, overweight, and weight gain over adult life are main determinants of elevated hs-CRP in a cohort of Mediterranean women. *Eur J Clin Nutr.* août 2010;64(8):873-8. doi:10.1038/ejcn.2010.69 PubMed PMID: 20517327.
193. Bastard JP, Maachi M, Lagathu C, Kim MJ, Caron M, Vidal H, et al. Recent advances in the relationship between obesity, inflammation, and insulin resistance. *Eur Cytokine Netw.* mars 2006;17(1):4-12. PubMed PMID: 16613757.
194. Hansen AB, Ravnskjær L, Loft S, Andersen KK, Bräuner EV, Baastrup R, et al. Long-term exposure to fine particulate matter and incidence of diabetes in the Danish Nurse Cohort. *Environ Int.* mai 2016;91:243-50. doi:10.1016/j.envint.2016.02.036 PubMed PMID: 26989812.
195. Imrich A, Ning Y, Lawrence J, Coull B, Gitin E, Knutson M, et al. Alveolar macrophage cytokine response to air pollution particles: oxidant mechanisms. *Toxicol Appl Pharmacol.* 1 févr 2007;218(3):256-64. doi:10.1016/j.taap.2006.11.033 PubMed PMID: 17222881; PubMed Central PMCID: PMC1852536.
196. Grabiec AM, Hussell T. The role of airway macrophages in apoptotic cell clearance following acute and chronic lung inflammation. *Semin Immunopathol.* juill 2016;38(4):409-23. doi:10.1007/s00281-016-0555-3 PubMed PMID: 26957481; PubMed Central PMCID: PMC4896990.
197. Robb CT, Regan KH, Dorward DA, Rossi AG. Key mechanisms governing resolution of lung inflammation. *Semin Immunopathol.* juill 2016;38(4):425-48. doi:10.1007/s00281-016-0560-6 PubMed PMID: 27116944; PubMed Central PMCID: PMC4896979.
198. Bolton J, Chaudhury S, MacGill RS, Early AM, King CR, Locke E, et al. Multiplex serological assay for establishing serological profiles of polymorphic, closely related peptide antigens. *MethodsX.* 2021;8:101345. doi:10.1016/j.mex.2021.101345 PubMed PMID: 34430249; PubMed Central PMCID: PMC8374401.
199. Bolton JS, Chaudhury S, Dutta S, Gregory S, Locke E, Pierson T, et al. Comparison of ELISA with electro-chemiluminescence technology for the qualitative and quantitative assessment of serological responses to vaccination. *Malaria Journal.* 17 avr 2020;19(1):159. doi:10.1186/s12936-020-03225-5
200. Celli BR, Locantore N, Yates J, Tal-Singer R, Miller BE, Bakke P, et al. Inflammatory biomarkers improve clinical prediction of mortality in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 15 mai 2012;185(10):1065-72. doi:10.1164/rccm.201110-1792OC PubMed PMID: 22427534.
201. Huang H, Huang X, Zeng K, Deng F, Lin C, Huang W. Interleukin-6 is a Strong Predictor of the Frequency of COPD Exacerbation Within 1 Year. *Int J Chron Obstruct Pulmon Dis.* 2021;16:2945-51. doi:10.2147/COPD.S332505 PubMed PMID: 34737559; PubMed Central PMCID: PMC8560075.

202. De Lorenzo A, Del Gobbo V, Premrov MG, Bigioni M, Galvano F, Di Renzo L. Normal-weight obese syndrome: early inflammation? *Am J Clin Nutr.* janv 2007;85(1):40-5. doi:10.1093/ajcn/85.1.40 PubMed PMID: 17209175.
203. Claudino WM, Quattrone A, Biganzoli L, Pestrin M, Bertini I, Di Leo A. Metabolomics: available results, current research projects in breast cancer, and future applications. *J Clin Oncol.* 1 juill 2007;25(19):2840-6. doi:10.1200/JCO.2006.09.7550 PubMed PMID: 17502626.
204. Bharucha T, Oeser C, Balloux F, Brown JR, Carbo EC, Charlett A, et al. STROBE-metagenomics: a STROBE extension statement to guide the reporting of metagenomics studies. *Lancet Infect Dis.* oct 2020;20(10):e251-60. doi:10.1016/S1473-3099(20)30199-7 PubMed PMID: 32768390; PubMed Central PMCID: PMC7406238.
205. BibliZap [Internet]. [cité 10 janv 2025]. Disponible sur: <https://biblizap.org/>
206. Rayyan [Internet]. [cité 5 juin 2025]. Disponible sur: https://rayyan.ai/users/sign_in
207. Higgins JPT, Morgan RL, Rooney AA, Taylor KW, Thayer KA, Silva RA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). *Environment International.* 1 avr 2024;186:108602. doi:10.1016/j.envint.2024.108602
208. Seow WJ, Low DY, Pan WC, Gunther SH, Sim X, Torta F, et al. Coffee, Black Tea, and Green Tea Consumption in Relation to Plasma Metabolites in an Asian Population. *Mol Nutr Food Res.* 29 oct 2020;e2000527. doi:10.1002/mnfr.202000527 PubMed PMID: 33120436.
209. van Roekel EH, Trijsburg L, Assi N, Carayol M, Achaintre D, Murphy N, et al. Circulating Metabolites Associated with Alcohol Intake in the European Prospective Investigation into Cancer and Nutrition Cohort. *Nutrients.* 22 mai 2018;10(5):654. doi:10.3390/nu10050654 PubMed PMID: 29789452; PubMed Central PMCID: PMC5986533.
210. Bagheri M, Willett W, Townsend MK, Kraft P, Ivey KL, Rimm EB, et al. A lipid-related metabolomic pattern of diet quality. *Am J Clin Nutr.* 10 déc 2020;112(6):1613-30. doi:10.1093/ajcn/nqaa242 PubMed PMID: 32936887; PubMed Central PMCID: PMC7727474.
211. Kim H, Hu EA, E Wong K, Yu B, Steffen LM, Seidemann SB, et al. Serum Metabolites Associated with Healthy Diets in African Americans and European Americans. *J Nutr.* 4 janv 2021;151(1):40-9. doi:10.1093/jn/nxaa338 PubMed PMID: 33244610; PubMed Central PMCID: PMC7779213.
212. Hernández-Alonso P, Becerra-Tomás N, Papandreou C, Bulló M, Guasch-Ferré M, Toledo E, et al. Plasma Metabolomics Profiles are Associated with the Amount and Source of Protein Intake: A Metabolomics Approach within the PREDIMED Study. *Mol Nutr Food Res.* juin 2020;64(12):e2000178. doi:10.1002/mnfr.202000178 PubMed PMID: 32378786; PubMed Central PMCID: PMC9245364.

213. Papandreou C, Hernández-Alonso P, Bulló M, Ruiz-Canela M, Yu E, Guasch-Ferré M, et al. Plasma Metabolites Associated with Coffee Consumption: A Metabolomic Approach within the PREDIMED Study. *Nutrients*. 8 mai 2019;11(5):1032. doi:10.3390/nu11051032 PubMed PMID: 31072000; PubMed Central PMCID: PMC6566346.
214. Hernández-Alonso P, Papandreou C, Bulló M, Ruiz-Canela M, Dennis C, Deik A, et al. Plasma Metabolites Associated with Frequent Red Wine Consumption: A Metabolomics Approach within the PREDIMED Study. *Mol Nutr Food Res*. sept 2019;63(17):e1900140. doi:10.1002/mnfr.201900140 PubMed PMID: 31291050; PubMed Central PMCID: PMC6771435.
215. Huang T, Zeleznik OA, Roberts AL, Balasubramanian R, Clish CB, Eliassen AH, et al. Plasma Metabolomic Signature of Early Abuse in Middle-Aged Women. *Psychosom Med*. 1 juin 2022;84(5):536-46. doi:10.1097/PSY.0000000000001088 PubMed PMID: 35471987; PubMed Central PMCID: PMC9167800.
216. Tabung FK, Liang L, Huang T, Balasubramanian R, Zhao Y, Chandler PD, et al. Identifying metabolomic profiles of inflammatory diets in postmenopausal women. *Clin Nutr*. mai 2020;39(5):1478-90. doi:10.1016/j.clnu.2019.06.010 PubMed PMID: 31255351; PubMed Central PMCID: PMC6918009.
217. Gu F, Derkach A, Freedman ND, Landi MT, Albanes D, Weinstein SJ, et al. Cigarette smoking behaviour and blood metabolomics. *Int J Epidemiol*. oct 2016;45(5):1421-32. doi:10.1093/ije/dyv330 PubMed PMID: 26721601; PubMed Central PMCID: PMC5100605.
218. Che X, Hong X, Gross S, Pearson C, Bartell T, Wang X, et al. Maternal Mediterranean-Style Diet Adherence during Pregnancy and Metabolomic Signature in Postpartum Plasma: Findings from the Boston Birth Cohort. *J Nutr*. mars 2024;154(3):846-55. doi:10.1016/j.tjnut.2024.01.022 PubMed PMID: 38278216; PubMed Central PMCID: PMC10942856.
219. Choi JJ, Kosciak RL, Jonaitis EM, Panyard DJ, Morrow AR, Johnson SC, et al. Assessing the Biological Mechanisms Linking Smoking Behavior and Cognitive Function: A Mediation Analysis of Untargeted Metabolomics. *Metabolites*. 16 nov 2023;13(11):1154. doi:10.3390/metabo13111154 PubMed PMID: 37999250; PubMed Central PMCID: PMC10673384.
220. Chen X, Zhao X, Jones MB, Harper A, de Seymour JV, Yang Y, et al. The relationship between hair metabolites, air pollution exposure and gestational diabetes mellitus: A longitudinal study from pre-conception to third trimester. *Front Endocrinol (Lausanne)*. 2022;13:1060309. doi:10.3389/fendo.2022.1060309 PubMed PMID: 36531491; PubMed Central PMCID: PMC9755849.
221. Zhang R, Sun X, Huang Z, Pan Y, Westbrook A, Li S, et al. Examination of serum metabolome altered by cigarette smoking identifies novel metabolites mediating smoking-BMI association. *Obesity (Silver Spring)*. avr 2022;30(4):943-52. doi:10.1002/oby.23386 PubMed PMID: 35258150; PubMed Central PMCID: PMC8957487.

222. Drouin-Chartier JP, Hernández-Alonso P, Guasch-Ferré M, Ruiz-Canela M, Li J, Wittenbecher C, et al. Dairy consumption, plasma metabolites, and risk of type 2 diabetes. *Am J Clin Nutr.* 1 juill 2021;114(1):163-74. doi:10.1093/ajcn/nqab047 PubMed PMID: 33742198; PubMed Central PMCID: PMC8246603.
223. Wang F, Baden MY, Guasch-Ferré M, Wittenbecher C, Li J, Li Y, et al. Plasma metabolite profiles related to plant-based diets and the risk of type 2 diabetes. *Diabetologia.* juill 2022;65(7):1119-32. doi:10.1007/s00125-022-05692-8 PubMed PMID: 35391539; PubMed Central PMCID: PMC9810389.
224. Cross AJ, Boca S, Freedman ND, Caporaso NE, Huang WY, Sinha R, et al. Metabolites of tobacco smoking and colorectal cancer risk. *Carcinogenesis.* juill 2014;35(7):1516-22. doi:10.1093/carcin/bgu071 PubMed PMID: 24648381; PubMed Central PMCID: PMC4076812.
225. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, et al. Identifying biomarkers of dietary patterns by using metabolomics. *Am J Clin Nutr.* févr 2017;105(2):450-65. doi:10.3945/ajcn.116.144501 PubMed PMID: 28031192; PubMed Central PMCID: PMC5267308.
226. Zheng Y, Yu B, Alexander D, Steffen LM, Nettleton JA, Boerwinkle E. Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study. *Am J Clin Nutr.* juin 2014;99(6):1470-8. doi:10.3945/ajcn.113.074070 PubMed PMID: 24760976; PubMed Central PMCID: PMC4021786.
227. Zheng Y, Yu B, Alexander D, Steffen LM, Boerwinkle E. Human metabolome associates with dietary intake habits among African Americans in the atherosclerosis risk in communities study. *Am J Epidemiol.* 15 juin 2014;179(12):1424-33. doi:10.1093/aje/kwu073 PubMed PMID: 24801555; PubMed Central PMCID: PMC4051875.
228. Salihovic S, Fall T, Ganna A, Broeckling CD, Prenni JE, Hyötyläinen T, et al. Identification of metabolic profiles associated with human exposure to perfluoroalkyl substances. *J Expo Sci Environ Epidemiol.* mars 2019;29(2):196-205. doi:10.1038/s41370-018-0060-y PubMed PMID: 30185940.
229. Cohen CC, Dabelea D, Michelotti G, Tang L, Shankar K, Goran MI, et al. Metabolome Alterations Linking Sugar-Sweetened Beverage Intake with Dyslipidemia in Youth: The Exploring Perinatal Outcomes among CHildren (EPOCH) Study. *Metabolites.* 17 juin 2022;12(6):559. doi:10.3390/metabo12060559 PubMed PMID: 35736491; PubMed Central PMCID: PMC9228193.
230. Walker ME, Song RJ, Xu X, Gerszten RE, Ngo D, Clish CB, et al. Proteomic and Metabolomic Correlates of Healthy Dietary Patterns: The Framingham Heart Study. *Nutrients.* 19 mai 2020;12(5):1476. doi:10.3390/nu12051476 PubMed PMID: 32438708; PubMed Central PMCID: PMC7284467.
231. Beuchel C, Becker S, Dittrich J, Kirsten H, Toenjes A, Stumvoll M, et al. Clinical and lifestyle related factors influencing whole blood metabolite levels - A comparative analysis

- of three large cohorts. *Mol Metab.* nov 2019;29:76-85. doi:10.1016/j.molmet.2019.08.010 PubMed PMID: 31668394; PubMed Central PMCID: PMC6734104.
232. Wang Y, Sha W, Wang H, Howard AG, Tsilimigras MCB, Zhang J, et al. Urbanization in China is associated with pronounced perturbation of plasma metabolites. *Metabolomics.* 19 sept 2020;16(10):103. doi:10.1007/s11306-020-01724-9 PubMed PMID: 32951074; PubMed Central PMCID: PMC7707273.
233. Gadgil MD, Kanaya AM, Sands C, Chekmeneva E, Lewis MR, Kandula NR, et al. Diet Patterns Are Associated with Circulating Metabolites and Lipid Profiles of South Asians in the United States. *J Nutr.* nov 2022;152(11):2358-66. doi:10.1093/jn/nxac191 PubMed PMID: 36774102; PubMed Central PMCID: PMC10157813.
234. Dai H, Leeder JS, Cui Y. A modified generalized Fisher method for combining probabilities from dependent tests. *Front Genet.* 2014;5:32. doi:10.3389/fgene.2014.00032 PubMed PMID: 24600471; PubMed Central PMCID: PMC3929847.
235. Tomofuji Y, Suzuki K, Kishikawa T, Shojima N, Hosoe J, Inagaki K, et al. Identification of serum metabolome signatures associated with retinal and renal complications of type 2 diabetes. *Commun Med.* 9 janv 2023;3(1):1. doi:10.1038/s43856-022-00231-3
236. Xia J, Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res.* juill 2010;38(Web Server issue):W71-77. doi:10.1093/nar/gkq329 PubMed PMID: 20457745; PubMed Central PMCID: PMC2896187.
237. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 25 oct 2005;102(43):15545-50. doi:10.1073/pnas.0506580102 PubMed PMID: 16199517; PubMed Central PMCID: PMC1239896.
238. Wieder C, Lai RPJ, Ebbels TMD. Single sample pathway analysis in metabolomics: performance evaluation and application. *BMC Bioinformatics.* 14 nov 2022;23(1):481. doi:10.1186/s12859-022-05005-1
239. Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* juill 2009;37(Web Server issue):W652-660. doi:10.1093/nar/gkp356 PubMed PMID: 19429898; PubMed Central PMCID: PMC2703878.
240. Benjamini Y, Heller R, Yekutieli D. Selective inference in complex research. *Philos Trans A Math Phys Eng Sci.* 13 nov 2009;367(1906):4255-71. doi:10.1098/rsta.2009.0127 PubMed PMID: 19805444; PubMed Central PMCID: PMC3263782.
241. 2001-18.pdf [Internet]. [cité 23 juill 2024]. Disponible sur: <https://storeylab.org/doc/2001-18.pdf>

242. Storey JD. A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2002;64(3):479-98. doi:10.1111/1467-9868.00346
243. He Y, Song P, Xu G. Adaptive bootstrap tests for composite null hypotheses in the mediation pathway analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 1 avr 2024;86(2):411-34. doi:10.1093/jrssb/qkad129
244. Song Y, Zhou X, Zhang M, Zhao W, Liu Y, Kardina SLR, et al. Bayesian shrinkage estimation of high dimensional causal mediation effects in omics studies. *Biometrics*. sept 2020;76(3):700-10. doi:10.1111/biom.13189 PubMed PMID: 31733066; PubMed Central PMCID: PMC7228845.
245. Gunzler D, Chen T, Wu P, Zhang H. Introduction to mediation analysis with structural equation modeling. *Shanghai Arch Psychiatry*. déc 2013;25(6):390-4. doi:10.3969/j.issn.1002-0829.2013.06.009 PubMed PMID: 24991183; PubMed Central PMCID: PMC4054581.
246. Carrico C, Gennings C, Wheeler DC, Factor-Litvak P. Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting. *J Agric Biol Environ Stat*. mars 2015;20(1):100-20. doi:10.1007/s13253-014-0180-3 PubMed PMID: 30505142; PubMed Central PMCID: PMC6261506.
247. Chadeau-Hyam M, Campanella G, Jombart T, Bottolo L, Portengen L, Vineis P, et al. Deciphering the complex: methodological overview of statistical models to derive OMICS-based biomarkers. *Environ Mol Mutagen*. août 2013;54(7):542-57. doi:10.1002/em.21797 PubMed PMID: 23918146.
248. Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 1 avr 2005;67(2):301-20. doi:10.1111/j.1467-9868.2005.00503.x
249. Warrack BM, Hnatyshyn S, Ott KH, Reily MD, Sanders M, Zhang H, et al. Normalization strategies for metabolomic analysis of urine samples. *J Chromatogr B Analyt Technol Biomed Life Sci*. 15 févr 2009;877(5-6):547-52. doi:10.1016/j.jchromb.2009.01.007 PubMed PMID: 19185549.
250. Nezami Ranjbar MR, Zhao Y, Tadesse MG, Wang Y, Ransom HW. Gaussian process regression model for normalization of LC-MS data using scan-level information. *Proteome Science*. 7 nov 2013;11(1):S13. doi:10.1186/1477-5956-11-S1-S13
251. Esko T, Hirschhorn JN, Feldman HA, Hsu YHH, Deik AA, Clish CB, et al. Metabolomic profiles as reliable biomarkers of dietary composition. *Am J Clin Nutr*. mars 2017;105(3):547-54. doi:10.3945/ajcn.116.144428 PubMed PMID: 28077380; PubMed Central PMCID: PMC5320413.
252. Abdi-Zscore2007-pretty.dvi [Internet]. [cité 19 oct 2025]. Disponible sur: <https://personal.utdallas.edu/~Herve/Abdi-Zscore2007-pretty.pdf>

253. Andraos S, Beck KL, Jones MB, Han TL, Conlon CA, de Seymour JV. Characterizing patterns of dietary exposure using metabolomic profiles of human biospecimens: a systematic review. *Nutr Rev.* 10 mars 2022;80(4):699-708. doi:10.1093/nutrit/nuab103 PubMed PMID: 35024860.
254. Fan S, Yeon A, Shahid M, Anger JT, Eilber KS, Fiehn O, et al. Sex-associated differences in baseline urinary metabolites of healthy adults. *Sci Rep.* 8 août 2018;8(1):11883. doi:10.1038/s41598-018-29592-3 PubMed PMID: 30089834; PubMed Central PMCID: PMC6082868.
255. Finnie MD, Ersser RS, Seakins JW, Snedden W. The occurrence and identification of o-hydroxyhippuric acid (salicyluric acid) in the urine of sick children. *Clin Chim Acta.* 1 juill 1976;70(1):171-8. doi:10.1016/0009-8981(76)90020-6 PubMed PMID: 947617.
256. VIDAL [Internet]. 2024 [cité 7 janv 2026]. KARDEGIC. Disponible sur: <https://www.vidal.fr/medicaments/gammes/kardegic-5146.html>
257. He Q, Chen YP, Li J, Wu H, Chen F, Li M, et al. Antioxidant and Photoprotective Activities of 3,4-Dihydroxybenzoic Acid and (+)-Catechin, Identified from *Schima argentea* Extract, in UVB-Irradiated HaCaT Cells. *Antioxidants.* févr 2025;14(2):241. doi:10.3390/antiox14020241
258. ResearchGate [Internet]. [cité 24 oct 2025]. (PDF) GCMS based metabolite profiling of first flush Darjeeling tea (orthodox) through the processing steps: First report. Disponible sur: https://www.researchgate.net/publication/362728062_GCMS_based_metabolite_profiling_of_first_flush_Darjeeling_tea_orthodox_through_the_processing_steps_First_report doi:10.21203/rs.3.rs-1775346/v1
259. Silva TD, Alves C, Oliveira H, Duarte IF. Biological Impact of Organic Extracts from Urban-Air Particulate Matter: An In Vitro Study of Cytotoxic and Metabolic Effects in Lung Cells. *International Journal of Molecular Sciences.* janv 2023;24(23):16896. doi:10.3390/ijms242316896
260. Lundberg I, Sollenberg J. Correlation of xylene exposure and methyl hippuric acid excretion in urine among paint industry workers. *Scand J Work Environ Health.* avr 1986;12(2):149-53. doi:10.5271/sjweh.2164 PubMed PMID: 3726497.
261. Wójcik P, Gęgotek A, Źarković N, Skrzydlewska E. Oxidative Stress and Lipid Mediators Modulate Immune Cell Functions in Autoimmune Diseases. *International Journal of Molecular Sciences.* janv 2021;22(2):723. doi:10.3390/ijms22020723
262. Kaźmierczak-Barańska J, Halczuk K, Karwowski BT. Thiamine (Vitamin B1)—An Essential Health Regulator. *Nutrients.* janv 2025;17(13):2206. doi:10.3390/nu17132206
263. Presnell CE, Bhatti G, Numan LS, Lerche M, Alkhateeb SK, Ghalib M, et al. Computational insights into the role of glutathione in oxidative stress. *Curr Neurovasc Res.* mai 2013;10(2):185-94. doi:10.2174/1567202611310020011 PubMed PMID: 23469953.

264. Wani AL, Ara A, Usmani JA. Lead toxicity: a review. *Interdiscip Toxicol.* juin 2015;8(2):55-64. doi:10.1515/intox-2015-0009 PubMed PMID: 27486361; PubMed Central PMCID: PMC4961898.
265. 0-Introduction-001/012.indd [Internet]. [cité 21 sept 2025]. Disponible sur: https://www.hauts-de-france.developpement-durable.gouv.fr/static/archive/site_drivre/environnement/IRE2007/IREen2006_web/Images/2-Air/Plomb.pdf
266. L'époque de l'essence au plomb est désormais révolue : 1,2 million de décès prématurés pourront être évités | ONU Info [Internet]. 2021 [cité 21 sept 2025]. Disponible sur: <https://news.un.org/fr/story/2021/08/1102732>
267. Korrick SA, Schwartz J, Tsaih SW, Hunter DJ, Aro A, Rosner B, et al. Correlates of Bone and Blood Lead Levels among Middle-aged and Elderly Women. *Am J Epidemiol.* 15 août 2002;156(4):335-43. doi:10.1093/aje/kwf042
268. Intoxication au plomb [Internet]. [cité 24 oct 2025]. Disponible sur: <https://www.who.int/fr/news-room/fact-sheets/detail/lead-poisoning-and-health>
269. Rubin's Pathology: Clinicopathologic Foundations of Medicine Rubin and Strayer. In: *Rubin's Pathology: Clinicopathologic Foundations of Medicine.* p. 266,267.
270. Rodríguez J, Mandalunis PM. A Review of Metal Exposure and Its Effects on Bone Health. *J Toxicol.* 2018;2018:4854152. doi:10.1155/2018/4854152 PubMed PMID: 30675155; PubMed Central PMCID: PMC6323513.
271. Yang B, Li X. Unveiling the Mechanisms of Bone Marrow Toxicity Induced by Lead Acetate Exposure. *Biol Trace Elem Res.* mars 2024;202(3):1041-66. doi:10.1007/s12011-023-03733-w PubMed PMID: 37378799.
272. <https://www.anses.fr/fr/system/files/VSR2013SA0042-1Ra.pdf> [Internet]. [cité 21 sept 2025]. Disponible sur: <https://www.anses.fr/fr/system/files/VSR2013SA0042-1Ra.pdf>
273. [VLEP2013SA0042.pdf](https://www.anses.fr/fr/system/files/VLEP2013SA0042.pdf) [Internet]. [cité 21 sept 2025]. Disponible sur: <https://www.anses.fr/fr/system/files/VLEP2013SA0042.pdf>
274. Niedzwiecki MM, Eggers S, Joshi A, Dolios G, Cantoral A, Lamadrid-Figueroa H, et al. Lead exposure and serum metabolite profiles in pregnant women in Mexico City. *Environ Health.* 10 déc 2021;20(1):125. doi:10.1186/s12940-021-00810-2 PubMed PMID: 34893088; PubMed Central PMCID: PMC8665540.
275. Eguchi A, Nomiya K, Sakurai K, Kim Trang PT, Viet PH, Takahashi S, et al. Alterations in urinary metabolomic profiles due to lead exposure from a lead-acid battery recycling site. *Environ Pollut.* nov 2018;242(Pt A):98-105. doi:10.1016/j.envpol.2018.06.071 PubMed PMID: 29966840.
276. Gao B, Chi L, Mahbub R, Bian X, Tu P, Ru H, et al. Multi-Omics Reveals that Lead Exposure Disturbs Gut Microbiome Development, Key Metabolites, and Metabolic

- Pathways. *Chem Res Toxicol*. 17 avr 2017;30(4):996-1005.
doi:10.1021/acs.chemrestox.6b00401 PubMed PMID: 28234468; PubMed Central PMCID: PMC5654721.
277. Ding Z, Kong Y, Shao X, Zhang Y, Ren C, Zhao X, et al. Growth, antioxidant capacity, intestinal morphology, and metabolomic responses of juvenile Oriental river prawn (*Macrobrachium nipponense*) to chronic lead exposure. *Chemosphere*. févr 2019;217:289-97. doi:10.1016/j.chemosphere.2018.11.034 PubMed PMID: 30419383.
278. Kelly RS, Bayne H, Spiro A, Vokonas P, Sparrow D, Weiss ST, et al. Metabolomic signatures of lead exposure in the VA Normative Aging Study. *Environ Res*. nov 2020;190:110022. doi:10.1016/j.envres.2020.110022 PubMed PMID: 32791250; PubMed Central PMCID: PMC7983049.
279. MetExplore [Internet]. [cité 15 oct 2025]. Disponible sur: <https://metexplore.toulouse.inrae.fr/metexplore2/>
280. MetExplore tutorial [Internet]. [cité 15 oct 2025]. Find identifiers. Disponible sur: <https://metexplore.pages.mia.inra.fr/metexplore-web/metexplore-training/docs/mappingmetabolomics/findids/>
281. ChEBI | Home [Internet]. [cité 15 oct 2025]. Disponible sur: <https://www.ebi.ac.uk/chebi/>
282. Kasperczyk S, Błaszczuk I, Dobrakowski M, Romuk E, Kapka-Skrzypczak L, Adamek M, et al. Exposure to lead affects male biothiols metabolism. *Ann Agric Environ Med*. 2013;20(4):721-5. PubMed PMID: 24364442.
283. Chrysant SG, Chrysant GS. The current status of homocysteine as a risk factor for cardiovascular disease: a mini review. *Expert Rev Cardiovasc Ther*. août 2018;16(8):559-65. doi:10.1080/14779072.2018.1497974 PubMed PMID: 29979619.
284. Navas-Acien A, Guallar E, Silbergeld EK, Rothenberg SJ. Lead exposure and cardiovascular disease--a systematic review. *Environ Health Perspect*. mars 2007;115(3):472-82. doi:10.1289/ehp.9785 PubMed PMID: 17431501; PubMed Central PMCID: PMC1849948.
285. Ledda C, Cannizzaro E, Lovreglio P, Vitale E, Stufano A, Montana A, et al. Exposure to Toxic Heavy Metals Can Influence Homocysteine Metabolism? *Antioxidants (Basel)*. 28 déc 2019;9(1):30. doi:10.3390/antiox9010030 PubMed PMID: 31905706; PubMed Central PMCID: PMC7022705.
286. Flora SJS, Mittal M, Mehta A. Heavy metal induced oxidative stress & its possible reversal by chelation therapy. *Indian J Med Res*. oct 2008;128(4):501-23. PubMed PMID: 19106443.
287. Al-Neamy FR, Almehdi AM, Alwash R, Pasha MA, Ibrahim A, Bener A. Occupational lead exposure and amino acid profiles and liver function tests in industrial workers. *Int J*

- Environ Health Res. juin 2001;11(2):181-8. doi:10.1080/09603120020047564 PubMed PMID: 11382350.
288. Ommati MM, Jamshidzadeh A, Heidari R, Sun Z, Zamiri MJ, Khodaei F, et al. Carnosine and Histidine Supplementation Blunt Lead-Induced Reproductive Toxicity through Antioxidative and Mitochondria-Dependent Mechanisms. *Biol Trace Elem Res.* janv 2019;187(1):151-62. doi:10.1007/s12011-018-1358-2 PubMed PMID: 29767280.
289. Wei J, Ji JS. Modification of vitamin B6 on the associations of blood lead levels and cardiovascular diseases in the US adults. *BMJ Nutr Prev Health.* déc 2020;3(2):180-7. doi:10.1136/bmjnph-2020-000088 PubMed PMID: 33521527; PubMed Central PMCID: PMC7841818.
290. Dobrakowski M, Pawlas N, Hudziec E, Kozłowska A, Mikołajczyk A, Birkner E, et al. Glutathione, glutathione-related enzymes, and oxidative stress in individuals with subacute occupational exposure to lead. *Environ Toxicol Pharmacol.* juill 2016;45:235-40. doi:10.1016/j.etap.2016.06.008 PubMed PMID: 27331344.
291. Hsu JM. Lead toxicity as related to glutathione metabolism. *J Nutr.* janv 1981;111(1):26-33. doi:10.1093/jn/111.1.26 PubMed PMID: 7452372.
292. Freeze HH, Boyce M, Zachara NE, Hart GW, Schnaar RL. Glycosylation Precursors. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., éditeurs. *Essentials of Glycobiology* [Internet]. 4th éd. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022 [cité 25 oct 2025]. Disponible sur: <http://www.ncbi.nlm.nih.gov/books/NBK579932/> PubMed PMID: 35536946.
293. Wang GY, Qin SL, Zheng YN, Geng HJ, Chen L, Yao JH, et al. Propionate promotes gluconeogenesis by regulating mechanistic target of rapamycin (mTOR) pathway in calf hepatocytes. *Animal Nutrition.* 1 déc 2023;15:88-98. doi:10.1016/j.aninu.2023.07.001
294. Wan H, Wang B, Cui Y, Wang Y, Zhang K, Chen C, et al. Low-level lead exposure promotes hepatic gluconeogenesis and contributes to the elevation of fasting glucose level. *Chemosphere.* août 2021;276:130111. doi:10.1016/j.chemosphere.2021.130111 PubMed PMID: 33691221.
295. Mostafalou S, Baeeri M, Bahadar H, Soltany-Rezaee-Rad M, Gholami M, Abdollahi M. Molecular mechanisms involved in lead induced disruption of hepatic and pancreatic glucose metabolism. *Environ Toxicol Pharmacol.* janv 2015;39(1):16-26. doi:10.1016/j.etap.2014.11.001 PubMed PMID: 25434758.
296. Bassuk SS, Rifai N, Ridker PM. High-sensitivity C-reactive protein: clinical importance. *Curr Probl Cardiol.* août 2004;29(8):439-93. PubMed PMID: 15258556.
297. Nothnagel M, Lu TT, Krawczak M. Hypotheses in genome-wide association scans. *Eur J Hum Genet.* oct 2008;16(10):1174-5. doi:10.1038/ejhg.2008.97

298. Rappaport SM. Biomarkers intersect with the exposome. *Biomarkers*. sept 2012;17(6):483-9. doi:10.3109/1354750X.2012.691553 PubMed PMID: 22672124; PubMed Central PMCID: PMC4763608.
299. Gabet S, Achour D, Anthérieu S, Chérot-Kornobis N, Crumeyrolle S, Dauchet L, et al. Innovative epidemiological design based on wearable microsensors to monitor mobility and related environmental exposures to better assess cardiovascular risks [Internet]. ISEE; août 2024 [cité 4 janv 2026]. Disponible sur: <https://lilloa.hal.science/hal-04768608>

VII. ANNEXES

Annexe 1 : Article “Association between residential air pollution exposure and cardiovascular risk factors in adults living in northern France” Environmental Geochemistry and Health

Association between residential air pollution exposure and lipid biomarkers in adults living in northern France

Marion Darras-Hostens[‡], Maxime Degrendel[‡], Philippe Amouyel[‡], Luc Dauchet^{‡,*}

[‡]Univ. Lille, INSERM, CHU Lille, Institut Pasteur de Lille, U1167 - RID-AGE - Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, F-59000 Lille, France

*Corresponding author at: Lille University Medical Center, Epidemiology Unit, 2 rue du Pr. Laguesse (MRRC), F-59037 Lille cedex, France

ABSTRACT

Background: Air pollution is associated with elevated cardiovascular mortality and an increase in cardiovascular risk factors. The literature data on associations between air pollution and blood lipid levels are contradictory. *Objectives:* To explore the relationship between residential exposure to atmospheric pollutants and lipid biomarker levels. *Methods:* We studied a sample of 2339 adult participants in the ELISABET study from the Dunkirk and Lille urban areas of France. The mean annual exposure to atmospheric pollutants (PM₁₀, NO₂ and SO₂) at the home address was estimated via an air dispersion model. The associations were probed in multivariate linear regression models. *Results:* The mean NO₂ level was 26.05 $\mu\text{g}/\text{m}^3$ in Lille and 19.96 $\mu\text{g}/\text{m}^3$ in Dunkirk. The mean PM₁₀ level was 27.02 $\mu\text{g}/\text{m}^3$ in Lille and 26.53 $\mu\text{g}/\text{m}^3$ in Dunkirk. We detected a significant association between exposure to air pollutants and the high-density lipoprotein (HDL) level: for a 2 $\mu\text{g}/\text{m}^3$ increment in PM₁₀, the HDL level decreased by 1.72% ($p=0.0037$). None of the associations with other lipid variables were significant. *Conclusion:* Long-term exposure of adults to moderate levels of ambient air pollution was associated with a decrement in HDL.

Keywords: lipids, air pollution, cross-sectional study, population-based study

1. Introduction

Exposure to particulate matter (PM) in the air increases the risk of cardiovascular events and cardiovascular death (1). Air pollution might also impact cardiovascular risk factors by altering lipid metabolism. Exposure to PM reduces the antioxidant capacity of high-density lipoproteins (HDLs) and thereby increases the size of atherosclerotic lesions (2,3). Chronic exposure to air pollutants causes oxidative stress, which is then responsible for changes in the endothelium, lipid metabolism, and atherosclerotic plaque progression (4). However, the epidemiological data on this topic differ from one study to another. In a systematic review, Gaio et al. (5) concluded that more evidence was needed because very few studies could be included in the meta-analysis. Other studies have since found associations with various markers of plasma cholesterol levels: HDLs in some studies (6–8), total cholesterol (TC) or low-density lipoproteins (LDL) in others (9,10), and all three variables in yet others (11,12). This heterogeneity suggests that the lipid profile is impacted by air pollution in an as-yet uncharacterized way.

The objective of the present study was to explore the association between mean residential levels of exposure to NO₂, PM₁₀ and SO₂ air pollutants on one hand and blood lipid levels (triglycerides (TGs), TC, LDLs, and HDLs) on the other.

2. Methods

2.1. Population

The data came from the *Enquête Littoral Souffle Air Biologie Environnement* (ELISABET) cross-sectional survey, which was based on a representative sample of men and women aged 40 to 65 living in the Lille and Dunkirk urban areas. Both areas have a high level of car traffic, and there is an industrial zone in the northwest part of the Dunkirk area. All participants were recruited between January 2011 and November 2013. The study methodology has been described in detail elsewhere (13–16). Participants were selected from electoral lists by random sampling, stratified on sex, age, and center (Lille or Dunkirk). Most data were collected at home via a detailed questionnaire. A blood sample was collected (at the participant's home, in most cases) by a trained, qualified nurse. Occasionally, the inclusion visit took place in a healthcare facility.

In line with the French legislation on biomedical research, the study protocol was approved by the local institutional review board (*CPP Nord Ouest IV*, Lille, France; reference: 2010-A00065-34; ClinicalTrials.gov identifier: NCT02490553). All the participants provided their written, informed consent.

In the present study, we selected ELISABET participants who had lived at their current address for at least the previous 12 months. Participants with missing data on pollutants, age, sex, center (Lille or Dunkirk), year of inclusion, educational level, smoking status, fruit and vegetable consumption, level of physical activity, alcohol consumption, body mass index, and diabetic status were excluded. We selected participants with a complete lipid profile: TGs,

LDL, HDL, and TC. Furthermore, we excluded participants taking lipid-lowering treatments. Lastly, participants who had not fasted for more than 10 hours at the time of blood sampling were excluded from our analysis.

For the analyses of LDL, we excluded participants with a TG level above 4 g/L; this enabled us to calculate the LDL level with the Friedewald formula ($LDL = TC - HDL - TG/5$ in g/L) (17).

2.2. Data on pollutants

Between 2010 and 2013 in Lille and between 2012 and 2013 in Dunkirk, the mean annual concentrations of NO₂, PM₁₀ and (for Dunkirk only) SO₂ were estimated by the ATMO Hauts-de-France air quality monitoring organization (Lille, France). ATMO uses an atmospheric dispersion modeling system that takes account of the topography, land use, pollutant emissions at ground level, meteorological data, and ambient pollution at the network's monitoring stations. Maps of the NO₂ and PM₁₀ concentrations in Lille and Dunkirk in 2015 can be viewed on the ATMO website (18).

We used a spatial resolution of 25 x 25 meters and evaluated the annual exposure of each residence within this grid. To this end, we averaged the values of the four corners of the grid, weighted by the inverse of the squared distance from each point. The overall exposure for each participant was then calculated as the average pollution level in the year of their inclusion in the study and in previous years. For the 468 participants included in 2011 in Dunkirk, we used 2012 pollution levels as the measure of exposure because pollutant concentration data for 2011 were not available (19).

2.3. Socioeconomic data on the neighborhoods

The population density for each neighborhood was calculated from data extracted from the *Ilots Regroupés pour l'Information Statistique* (IRIS) submunicipal area databases, which are curated by the French National Institute for Statistics and Economic Studies (INSEE: www.insee.fr). The IRIS unit is the smallest census unit available in France. The social deprivation index at the IRIS unit level was used to take account of the participants' socioeconomic level.

2.4. Lipid biomarkers

TC and TGs were measured using standard enzyme assays. HDL was measured after the precipitation of apolipoprotein B with phosphotungstate-magnesium chloride. As mentioned above, the LDL level was calculated according to the Friedewald formula, and the patient had fasted for 10 hours before providing the blood sample for the lipid assays.

From January to March 2010, blood assays were performed in the Pasteur Institute of Lille's genomic analysis laboratory and Lille University Medical Center's central laboratory. From April 2010 to November 2013, all blood tests were performed in the central laboratory only.

2.5. Statistical analyses

Associations between exposure to each air pollutant (PM₁₀, SO₂, and NO₂) and the dependent variables (blood lipid levels) were estimated using multiple linear regression models. In order to normalize the data distribution, all dependent variables were log-transformed. Hence, the regression coefficients were expressed as the percentage change [95% confidence interval (CI)] per 2 µg/m³ of PM₁₀ and SO₂ or per 5 µg/m³ of NO₂. These increments for residential exposure were chosen because there were of the same order of magnitude as the interquartile range of the measurements in Lille and Dunkirk. The values were back-transformed and converted into a percentage change by multiplying by 100 and then subtracting 100 percentage points.

For each outcome and pollutant, we fitted generalized linear regression models to potential confounders. Thus, model 1 was adjusted for age, sex, center (Lille or Dunkirk), and year of inclusion. For model 2, we added the educational level, smoking status, daily fruit and vegetable consumption, the level of physical activity, and alcohol consumption to the variables in model 1. For the fully adjusted model 3, we added BMI and diabetic status to the variables in model 2. To assess potential changes in interaction effects, we tested interactions between each pollutant and center (Lille or Dunkirk), age, sex, BMI, and smoking status. Furthermore, we hypothesized that the characteristics of the living area were associated with pollution and with changes in lipid status. We therefore performed sensitivity analyses by adding the social deprivation index or the population density (models 4 and 5, respectively). The threshold for statistical significance was set to $p < 0.05$. Statistical analyses were performed using R software (version 4.3.0).

3. Results

3.1. Population

There were 3276 participants in the ELISABET survey. Of these, 101 were excluded because they withdrew their consent to the use of personal data. In the analyses of TGs, TC, and HDL, 43 participants were excluded (leaving 3132 participants) because they had lived for less than 12 months at the reported home address. Next, we excluded 214 participants because they lacked data on covariates. Of the 2918 remaining participants, 17 were excluded because they had missing lipid data, and 466 were excluded because they were on lipid-lowering therapy. This left 2435 participants. After the exclusion of 96 participants because they had not fasted prior to blood sample collection, the final sample was composed of 2339 participants. For the LDL analysis, we excluded 18 participants who had missing data on LDL and 12 participants who had TG levels above 4 g/L (preventing use of the Friedewald formula (25)). We therefore included 2309 participants in the LDL analysis.

The characteristics of the participants in the full analysis set are summarized in Table 1. The mean (standard deviation (SD)) age was 52.26 (7.07) in Lille and 52.26 (7.18) in Dunkirk. The proportion of men was 44.20% (n=554) in Lille and 46.70% (n=507) in Dunkirk. Compared with participants in Lille, participants in Dunkirk had a higher BMI and greater prevalences of diabetes and alcohol consumption but a lower educational level and lower levels of physical activity and daily fruit and vegetable consumption.

3.2. Pollution

The concentrations of PM₁₀ and NO₂ were higher in Lille than Dunkirk (27.02 $\mu\text{g}/\text{m}^3$ vs. 26.53 $\mu\text{g}/\text{m}^3$ for PM₁₀, and 26.05 $\mu\text{g}/\text{m}^3$ vs. 19.96 $\mu\text{g}/\text{m}^3$ for NO₂, respectively; Table 1). The mean exposure level for SO₂ (measured in Dunkirk only) was 3.28 $\mu\text{g}/\text{m}^3$. Pearson's coefficient for the correlation between NO₂ and PM₁₀ was 0.87 in Lille and 0.83 in Dunkirk. The coefficient was 0.64 for the correlation between SO₂ and NO₂ and 0.72 for the correlation between SO₂ and PM₁₀.

3.3. Associations between residential air pollution and lipid biomarkers

Multivariate analyses showed a significant association between PM₁₀ exposure and HDL concentration (Table 2). Thus, a 2 $\mu\text{g}/\text{m}^3$ increment in PM₁₀ was significantly associated with a 1.72% decrement in the HDL level (p=0.0037). The HDL concentration rose significantly with exposure to NO₂ and SO₂ but only in model 1 (with decrements of 1.14% (p=0.0461) and 2.14% (p=0.0097), respectively).

In Dunkirk, the association between SO₂ exposure and the TG level found in model 1 was not significant in the full model (model 3: p=0.162).

3.4. Sensitivity analyses

In the sensitivity analyses (Table 3), the association between PM₁₀ exposure and HDL level was still significant in model 4 (i.e. model 3 adjusted for a social deprivation index; (p=0.023). The interactions between the urban area, BMI, smoking status, age, sex, and pollution levels were not statistically significant.

4. Discussion

Our results showed that a decrement in blood HDL levels was associated with long-term exposure to PM₁₀. For an increment of 2 $\mu\text{g}/\text{m}^3$ PM₁₀, we observed a 1.72% decrement in the HDL level in model 3 (p=0.0037).

Interventional studies have found that short-term exposure to PM₁₀ leads to a decrease in HDL levels (20,21). However, studies of long-term exposure have given contradictory results (22–26). As in the present research, a cross-sectional study conducted in southern Israel found a 1.29% decrement in HDL for a 20 $\mu\text{g}/\text{m}^3$ increment in PM₁₀ (27). A cross-sectional study in three large Chinese cities found 1.4% and 1.1% decrements in HDL for a 10 $\mu\text{g}/\text{m}^3$ increase in PM₁₀ and PM_{2.5} (28). Conversely, Shanley et al. and Chuang et al. did not find an association (22,26). These discrepancies remain to be explained. They might be related to the characteristics of the respective study populations and pollutant exposures (duration, level, and source) and/or differences in the measurement methods.

The association between air pollution and the HDL level has a possible biological explanation. Exposure to air pollutants triggers an inflammation signaling cascade - notably through the synthesis of serum amyloid A (SAA). The latter replaces apolipoprotein ApoA1, which serves as a component of HDL lipoprotein and a cofactor for lecithin-cholesterol acyltransferase (an enzyme that esterifies cholesterol prior to its transport into the liver). Once in the liver, cholesterol is partly recycled and eliminated via the enterohepatic cycle. Furthermore, SAA-HDL facilitates cholesterol efflux into the macrophages of the vascular intima; this prompts differentiation into foam cells and accelerates the formation of atherosclerotic plaques (29,30).

The reduction in the strength of the pollution-HDL association after adjustment for neighborhood characteristics in the sensitivity analysis suggests that part of this association is due to confounding geographic factors. Nevertheless, neighborhood characteristics might have a direct or indirect causal association with air pollution. This is particularly true for population density. The decrease in the strength of the association might therefore also be explained by overfitting.

Strengths and weaknesses

Our study had some limitations. Firstly, air pollutant exposure was modeled at the subject's home only; hence, occupational exposure and other exposure during activities outside the home were not taken into account. Secondly, and even though we adjusted for many covariables, the association between the air pollution level and the HDL level might have been biased by unmeasured confounders (e.g. inflammation markers like SAA). Thirdly, we

measured lipid levels in blood samples collected after the participant had fasted for at least 10 hours, rather than the 12 hours recommended in the guidelines. Lastly, we did not correct for multiple comparisons because the statistical tests were not independent. Nevertheless, the association between HDL and PM₁₀ remained statistically significant after correction for 12 tests (i.e. four lipid markers with three air pollutants).

The study also had some strengths. Firstly, we used standardized measurement methods and protocols to analyzed a general population sample. Secondly, exposure to pollutants was measured with a protocol used routinely to elaborate air quality indices. Thirdly, exposure to pollutants at the participant's home was modeled at a high spatial resolution.

5. Conclusion

The results of our cross-sectional study of a middle-aged urban population highlighted an association between long-term exposure to PM₁₀ and low HDL levels. This finding suggests that ambient air pollution is a risk factor for changes in HDL metabolism. Further longitudinal studies are needed to investigate the putative causal nature of this association.

Acknowledgements

We thank Lille University Hospital (especially the Centre de Biologie et de Pathologie), the University of Lille, the Institut Pasteur of Lille (especially the Departments of Médecine Préventive, Biologie Spécialisée and Médecine du Travail, and the Laboratoire d'Analyses Génomiques) and the Centre Hospitalier Général de Dunkerque (especially the Departments of Biology and Pneumology). We particularly thank the nurses, physicians and secretarial staff at the University of Lille and the Institut Pasteur of Lille. Lastly, we thank the French Ministère de l'Enseignement Supérieur et de la Recherche, the Hauts de France Region and the European Regional Development Fund for financial support.

Funding

This work was funded by Lille University Medical Center (CHU de Lille, Lille, France), the Nord Pas-de-Calais Regional Council, and the European Regional Development Fund (ERDF-FEDER Presage N°36034) as part of the CPER Institut de Recherche en ENvironnement Industriel (IRENI) program. The work was also funded by the French government through the Programme Investissement d'Avenir (I-SITE ULNE / ANR-16-IDEX-0004 ULNE), managed by the Agence Nationale de la Recherche as part of the project "Santé Environnement: du risque territorial au risque individuel" (also funded by Lille European Metropole). Lastly, this work is part of the CPER's CLIMIBIO research project.

References

1. Brook RD, Rajagopalan S, Pope CA, Brook JR, Bhatnagar A, Diez-Roux AV, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*. 2010 Jun 1;121(21):2331–78.
2. Li R, Navab M, Pakbin P, Ning Z, Navab K, Hough G, et al. Ambient ultrafine particles alter lipid metabolism and HDL anti-oxidant capacity in LDLR-null mice. *Journal of Lipid Research*. 2013 Jun;54(6):1608–15.
3. Rao X, Zhong J, Maiseyeu A, Gopalakrishnan B, Villamena FA, Chen LC, et al. CD36-Dependent 7-Ketocholesterol Accumulation in Macrophages Mediates Progression of

Atherosclerosis in Response to Chronic Air Pollution Exposure: Novelty and Significance. *Circulation Research*. 2014 Oct 10;115(9):770–80.

4. Bai Y, Sun Q. Fine particulate matter air pollution and atherosclerosis: Mechanistic insights. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2016 Dec;1860(12):2863–8.
5. Gaio V, Roquette R, Dias CM, Nunes B. Ambient air pollution and lipid profile: Systematic review and meta-analysis. *Environ Pollut*. 2019 Nov;254(Pt B):113036.
6. Kim KN, Ha B, Seog W, Hwang IU. Long-term exposure to air pollution and the blood lipid levels of healthy young men. *Environ Int*. 2022 Mar;161:107119.
7. Wu XM, Broadwin R, Basu R, Malig B, Ebisu K, Gold EB, et al. Associations between fine particulate matter and changes in lipids/lipoproteins among midlife women. *Sci Total Environ*. 2019 Mar 1;654:1179–86.
8. He ZZ, Guo PY, Xu SL, Zhou Y, Jalaludin B, Leskinen A, et al. Associations of Particulate Matter Sizes and Chemical Constituents with Blood Lipids: A Panel Study in Guangzhou, China. *Environ Sci Technol*. 2021 Apr 20;55(8):5065–75.
9. Gui ZH, Yang BY, Zou ZY, Ma J, Jing J, Wang HJ, et al. Exposure to ambient air pollution and blood lipids in children and adolescents: A national population based study in China. *Environ Pollut*. 2020 Nov;266(Pt 3):115422.
10. Li J, Yao Y, Xie W, Wang B, Guan T, Han Y, et al. Association of long-term exposure to PM_{2.5} with blood lipids in the Chinese population: Findings from a longitudinal quasi-experiment. *Environ Int*. 2021 Jun;151:106454.
11. McGuinn LA, Schneider A, McGarrah RW, Ward-Caviness C, Neas LM, Di Q, et al. Association of long-term PM_{2.5} exposure with traditional and novel lipid measures related to cardiovascular disease risk. *Environ Int*. 2019 Jan;122:193–200.
12. Wang L, Chen G, Pan Y, Xia J, Chen L, Zhang X, et al. Association of long-term exposure to ambient air pollutants with blood lipids in Chinese adults: The China Multi-Ethnic Cohort study. *Environ Res*. 2021 Jun;197:111174.
13. Riant M, Meirhaeghe A, Giovannelli J, Occelli F, Havet A, Cuny D, et al. Associations between long-term exposure to air pollution, glycosylated hemoglobin, fasting blood glucose and diabetes mellitus in northern France. *Environ Int*. 2018 Nov;120:121–9.
14. Quach A, Giovannelli J, Chérot-Kornobis N, Ciuchete A, Clément G, Matran R, et al. Prevalence and underdiagnosis of airway obstruction among middle-aged adults in northern France: The ELISABET study 2011-2013. *Respir Med*. 2015 Dec;109(12):1553–61.
15. Clément G, Giovannelli J, Cottel D, Montaye M, Ciuchete A, Dallongeville J, et al. Changes over time in the prevalence and treatment of cardiovascular risk factors, and

contributions to time trends in coronary mortality over 25 years in the Lille urban area (northern France). *Archives of Cardiovascular Diseases*. 2017 Dec;110(12):689–99.

16. Giovannelli J, Trouiller P, Hulo S, Chérot-Kornobis N, Ciuchete A, Edmé JL, et al. Low-grade systemic inflammation: a partial mediator of the relationship between diabetes and lung function. *Ann Epidemiol*. 2018 Jan;28(1):26–32.

17. Annales du Contrôle National de Qualité des Analyses de Biologie Médicale [Internet]. [cited 2018 Aug 30]. Available from: https://ansm.sante.fr/var/ansm_site/storage/original/application/cf2f308ef14ebd9ca18fc7041d844655.pdf

18. RAPENEAU P, PATRIS J. COMMUNAUTÉ URBAINE D'ARRAS. 2014;8.

19. Miller KA, Siscovick DS, Sheppard L, Shepherd K, Sullivan JH, Anderson GL, et al. Long-Term Exposure to Air Pollution and Incidence of Cardiovascular Events in Women. *New England Journal of Medicine*. 2007 Feb 1;356(5):447–58.

20. Tomao E, Tiziana PB, Rosati V, Marcellini L, Tomei F. The effects of air pollution on the lipid balance of traffic police personnel. *Ann Saudi Med*. 2002 Nov;22(5–6):287–90.

21. Huang YCT, Rappold AG, Graff DW, Ghio AJ, Devlin RB. Synergistic effects of exposure to concentrated ambient fine pollution particles and nitrogen dioxide in humans. *Inhalation Toxicology*. 2012 Oct;24(12):790–7.

22. Shanley RP, Hayes RB, Cromar KR, Ito K, Gordon T, Ahn J. Particulate Air Pollution and Clinical Cardiovascular Disease Risk Factors. *Epidemiology*. 2016 Mar;27(2):291–8.

23. Chen Z, Salam MT, Toledo-Corral C, Watanabe RM, Xiang AH, Buchanan TA, et al. Ambient Air Pollutants Have Adverse Effects on Insulin and Glucose Homeostasis in Mexican Americans. *Diabetes Care*. 2016 Apr;39(4):547–54.

24. Bell G, Mora S, Greenland P, Tsai M, Gill E, Kaufman JD. Association of Air Pollution Exposures With High-Density Lipoprotein Cholesterol and Particle Number: The Multi-Ethnic Study of Atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2017 May;37(5):976–82.

25. Cai Y, Hansell AL, Blangiardo M, Burton PR, BioSHaRE, de Hoogh K, et al. Long-term exposure to road traffic noise, ambient air pollution, and cardiovascular risk factors in the HUNT and lifelines cohorts. *Eur Heart J*. 2017 Aug 1;38(29):2290–6.

26. Chuang KJ, Yan YH, Chiu SY, Cheng TJ. Long-term air pollution exposure and risk factors for cardiovascular diseases among the elderly in Taiwan. *Occup Environ Med*. 2011 Jan;68(1):64–8.

27. Yitshak Sade M, Kloog I, Liberty IF, Schwartz J, Novack V. The Association Between Air Pollution Exposure and Glucose and Lipids Levels. *The Journal of Clinical Endocrinology & Metabolism*. 2016 Jun;101(6):2460–7.

28. Yang BY, Bloom MS, Markevych I, Qian ZM, Vaughn MG, Cummings-Vaughn LA, et al. Exposure to ambient air pollution and blood lipids in adults: The 33 Communities Chinese Health Study. *Environ Int.* 2018 Oct;119:485–92.
29. Saber AT, Jacobsen NR, Jackson P, Poulsen SS, Kyjovska ZO, Halappanavar S, et al. Particle-induced pulmonary acute phase response may be the causal link between particle inhalation and cardiovascular disease: Particle inhalation and cardiovascular disease. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology.* 2014 Nov;6(6):517–31.
30. Bourdon JA, Halappanavar S, Saber AT, Jacobsen NR, Williams A, Wallin H, et al. Hepatic and Pulmonary Toxicogenomic Profiles in Mice Intratracheally Instilled With Carbon Black Nanoparticles Reveal Pulmonary Inflammation, Acute Phase Response, and Alterations in Lipid Homeostasis. *Toxicol Sci.* 2012 Jun;127(2):474–84.

Table 1: Characteristics of the study participants

	Lille	Dunkirk
n (%)	1254 (53.61)	1085 (46.39)
Sex (men), n (%)	554 (44.20)	507 (46.70)
Age (years), mean (SD)	52.26 (7.07)	52.26 (7.18)
Smokers, n (%)	242 (19.30)	196 (18.10)
BMI (kg/m ²), mean (SD)	26.11 (4.88)	26.94 (4.90)
Educational level, n (%)		
5 or more years of higher education	320 (25.50)	157 (14.50)
1 to 4 years of higher education	280 (22.30)	176 (16.20)
no higher education	570 (45.50)	640 (59.00)
no secondary education	84 (6.70)	112 (10.30)
Physical activity level (MET), mean (SD)	3464 (4528)	2058 (3041)
Diabetes, n (%)	96 (7.70)	109 (10.00)
Fruit and vegetable consumption (per days), mean (SD)	2.53 (1.34)	2.33 (1.30)
Alcohol consumption (g/days), mean (SD)	8.83 (13.48)	10.24 (15.77)
Lipid profile (g/L), mean (SD)		
HDL	0.60 (0.15)	0.58 (0.15)
Triglycerides	1.12 (0.88)	1.15 (0.79)
Total cholesterol	2.25 (0.40)	2.25 (0.37)
LDL	1.43 (0.34)	1.45 (0.32)
Pollutants (µg/m ³), mean (SD)		
SO ₂		3.28 (1.62)
PM ₁₀	27.02 (1.98)	26.53 (1.00)
NO ₂	26.05 (5.31)	19.96 (2.19)

Abbreviations: SD = standard deviation; BMI = body mass index; HDL = high density lipoprotein; LDL = low density lipoprotein; PM₁₀ = particulate matter with an aerodynamic diameter below 10 µm; NO₂ = nitrogen dioxide; SO₂ = sulfur dioxide

Table 2: Associations between exposure to atmospheric pollutants and serum biomarkers (g/L)

Biomarker	NO ₂			PM ₁₀			SO ₂		
	Number	Percentage change for 5 µg/m ³ [95%CI]	p trend	Number	Percentage change for 2 µg/m ³ [95%CI]	p-value	Number	Percentage change for 2 µg/m ³ [95%CI]	p-value
Triglycerides									
Model 1	2339	0.46 [-1.79; 2.77]	0,690	2339	2.36 [-0.17; 4.96]	0,068	1085	4.48 [1.03; 8.04]	0,0107
Model 2	2339	0.10 [-2.16; 2.41]	0,931	2339	1.62 [-0.91; 4.21]	0,213	1085	3.31 [-0.10; 6.84]	0,0573
Model 3	2339	0.28 [-1.87; 2.47]	0,801	2339	1.38 [-1.01; 3.84]	0,260	1085	2.30 [-0.90; 5.62]	0,162
Total cholesterol									
Model 1	2339	-0.46 [-1.27; 0.37]	0,279	2339	-0.41 [-1.31; 0.51]	0,383	1085	0.63 [-0.59; 1.85]	0,314
Model 2	2339	-0.28 [-1.11; 0.56]	0,517	2339	-0.23 [-1.15; 0.69]	0,623	1085	0.44 [-0.77; 1.68]	0,477
Model 3	2339	-0.24 [-1.07; 0.60]	0,572	2339	-0.20 [-1.12; 0.72]	0,665	1085	0.53 [-0.69; 1.76]	0,396
HDL									
Model 1	2339	-1.14 [-2.25; -0.020]	0,0461	2339	-2.35 [-3.56; -1.13]	< 0.001	1085	-2.14 [-3.73; -0.53]	0,0097
Model 2	2339	-0.76 [-1.87; 0.35]	0,179	2339	-1.81 [-3.01; -0.59]	0,0038	1085	-1.92 [-3.48; -0.33]	0,0182
Model 3	2339	-0.87 [-1.92; 0.18]	0,105	2339	-1.72 [-2.86; -0.56]	0,0037	1085	-1.45 [-2.94; 0.060]	0,0607
LDL									
Model 1	2309	-0.65 [-1.79; 0.50]	0,266	2309	-0.47 [-1.73; 0.80]	0,466	1068	1.22 [-0.46; 2.92]	0,155
Model 2	2309	-0.47 [-1.63; 0.70]	0,426	2309	-0.31 [-1.59; 0.99]	0,637	1068	1.05 [-0.65; 2.77]	0,229
Model 3	2309	-0.40 [-1.55; 0.77]	0,502	2309	-0.27 [-1.55; 1.02]	0,681	1068	1.10 [-0.59; 2.82]	0,205

Abbreviations: NO₂ = nitrogen dioxide; PM₁₀ = particulate matter with an aerodynamic diameter below 10 µm; SO₂ = sulfur dioxide; HDL = high-density lipoprotein; LDL = low-density lipoprotein

Model 1: adjustment for age, sex, center, and year of inclusion

Model 2: adjustment for age, sex, center, year of inclusion, educational level, smoking status, fruit and vegetable consumption, physical activity level, and alcohol consumption

Model 3: adjustment for age, sex, center, year of inclusion, educational level, smoking status, fruit and vegetable consumption, physical activity level, alcohol consumption, body mass index, and diabetic status

Table 3: Associations between exposure to atmospheric pollutants and serum biomarkers (g/L) - sensitivity analyses

Biomarker	NO ₂			PM ₁₀			SO ₂		
	Number	Percentage change for 5 µg/m ³ [95%CI]	p trend	Number	Percentage change for 2 µg/m ³ [95%CI]	p-value	Number	Percentage change for 2 µg/m ³ [95%CI]	p-value
Triglycerides									
Model 4	2339	0.21 [-1.94; 2.40]	0,849	2339	0.86 [-1.56; 3.33]	0,490	1085	1.60 [-1.91; 5.23]	0,376
Model 5	2339	0.45 [-1.94; 2.91]	0,713	2339	1.27 [-1.42; 4.03]	0,359	1085	1.81 [-1.77; 5.52]	0,326
Total cholesterol									
Model 4	2339	-0.24 [-1.07; 0.60]	0,577	2339	-0.19 [-1.12; 0.75]	0,696	1085	0.71 [-0.64; 2.07]	0,303
Model 5	2339	0.09 [-0.84; 1.02]	0,857	2339	0.18 [-0.85; 1.23]	0,729	1085	0.98 [-0.39; 2.37]	0,162
HDL									
Model 4	2339	-0.82 [-1.86; 0.23]	0,125	2339	-1.36 [-2.52; -0.19]	0,023	1085	-0.80 [-2.45; 0.88]	0,351
Model 5	2339	-0.12 [-1.28; 1.07]	0,847	2339	-0.70 [-1.99; 0.61]	0,293	1085	-0.36 [-2.05; 1.36]	0,680
LDL									
Model 4	2309	-0.40 [-1.56; 0.77]	0,498	2309	-0.32 [-1.61; 1.00]	0,636	1068	1.09 [-0.77; 3.00]	0,253
Model 5	2309	-0.22 [-1.51; 1.09]	0,741	2309	-0.08 [-1.52; 1.38]	0,91	1068	1.34 [-0.57; 3.28]	0,172

Abbreviations: NO₂ = nitrogen dioxide; PM₁₀ = particulate matter with an aerodynamic diameter below 10 µm; SO₂ = sulfur dioxide; HDL = high-density lipoprotein; LDL = low-density lipoprotein

Model 4: adjustment for age, sex, center, year of inclusion, educational level, smoking status, fruit and vegetable consumption, physical activity, alcohol consumption, body mass index, diabetic status, and social deprivation index

Model 5: adjustment for age, sex, center, year of inclusion, educational level, smoking status, fruit and vegetable consumption, physical activity, alcohol consumption, body mass index, diabetic status, social deprivation index, and population density

Annexe 2 : “Article Short-term and residential exposure to air pollution: Associations with inflammatory biomarker levels in adults living in northern France” Science of the Total Environment

Short-term and residential exposure to air pollution: associations with inflammatory biomarker levels in adults living in northern France

Authors

Marion Darras-Hostens^a, Djamel Achour^b, Manon Muntaner^a, Céline Grare^b, Gianni Zarcone^b, Guillaume Garçon^b, Philippe Amouyel^a, Farid Zerimech², Régis Matran^b, Jean-Marc Lo Guidice^b, Luc Dauchet^a

^a Univ. Lille, INSERM, CHU Lille, Institut Pasteur de Lille, U1167 - RID-AGE - Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, F-59000 Lille, France

^b Univ. Lille, CHU Lille, Institut Pasteur de Lille, ULR 4483 – IMPECS – IMPact de l'Environnement Chimique sur la Santé, F-59000 Lille, France

E-mail addresses: marion.hostens.etu@univ-lille.fr (M. Darras-Hostens), djamel.achour@univ-lille.fr (D. Achour), manon.muntaner@pasteur-lille.fr (M. Muntaner), celine.grare@univ-lille.fr (C. Grare), gianni.zarcone@univ-lille.fr (G. Zarcone), guillaume.garcon@univ-lille.fr (G. Garçon), philippe.amouyel@univ-lille.fr (P. Amouyel), Farid.ZERIMECH@chu-lille.fr (F. Zerimech), regis.matran@univ-lille.fr (R. Matran), jean-marc.lo-guidice@univ-lille.fr (J.M. Lo Guidice), luc.dauchet@chru-lille.fr (L. Dauchet)

*Corresponding author at: University Hospital of Lille, Epidemiology Unit; 2, rue du Pr. Laguesse (MRRC), F-59037 Lille Cedex, France; Tel: +33 3 20 44 55 18, fax: +33 3 20 44 69 45, E-mail: luc.dauchet@chru-lille.fr

HIGHLIGHTS

- Low-grade inflammation might explain air pollution's harmful effects on health
- We assessed associations between air pollution and inflammatory biomarker levels
- Short-term exposure was associated with higher C-reactive protein levels
- The associations with plasma cytokine levels were not statistically significant
- Repeated blood sampling is needed to better characterize these associations

ABSTRACT

Introduction: Air pollution has an impact on health, and low-grade inflammation might be one of the underlying mechanisms. The objective of the present study of adults from northern France was to assess the associations between short-term and residential exposure to air pollution and levels of various inflammatory biomarkers.

Methods: The cross-sectional *Enquête Littoral Souffle Air Biologie Environnement* (ELISABET) study was conducted from 2011 to 2013 in the Lille and Dunkirk urban areas of northern France. Here, we evaluated the associations between PM₁₀, NO₂ and O₃ exposure (on the day of the blood sample collection and on the day before, and the mean annual residential level) and levels of the inflammatory biomarkers high-sensitivity C-reactive protein (hsCRP), interleukin (IL)-1 β , IL-6, IL-8, IL-10, IL-17A, IL-22, and tumor necrosis factor α .

Results: We assessed 3074 participants for the association with hsCRP and a subsample of 982 non-smokers from Lille for the association with plasma cytokine levels. A 10 $\mu\text{g}/\text{m}^3$ increment in PM₁₀ and NO₂ levels on the day of sample collection and on the day before was associated with a higher hsCRP concentration (3.43% [0.68; 6.25] and 1.75% [-1.96; 5.61], respectively, whereas a 10 $\mu\text{g}/\text{m}^3$ increment in O₃ was associated with lower hsCRP concentration (-1.2% [-3.95; 1.64]). The associations between mean annual exposure and the hsCRP level were not significant. Likewise, the associations between exposure and plasma cytokine levels were not statistically significant.

Conclusion: Short-term exposure to air pollution was associated with higher serum hsCRP levels in adult residents of two urban areas in northern France. Our results suggest that along with other factors, low-grade inflammation might explain the harmful effects of air pollution on health.

1. Introduction

Air pollution has a major impact on health: according to the World Health Organization, it causes the death of 7 million people per year (1). The effects of air pollution on the respiratory and cardiovascular systems are particularly harmful (2). Low-grade, systemic inflammation is not caused by a specific chronic or acute disease but promotes the onset and/or progression of certain chronic diseases, including cardiovascular disease(3), diabetes (4), cancer (5), and even depression (6). Thus, low-grade pulmonary and systemic inflammation might increase the potential for coagulation, thrombosis and, ultimately, cardiovascular and ischemic events (7).

Furthermore, particulate matter (PM) and some of its constituents have pro-inflammatory effects on the body. PM might increase low-grade inflammation and thus the above-mentioned health problems. More specifically, the biological changes induced by PM exposure may occur through several mechanistic pathways (2,8): (i) the release of pro-inflammatory mediators (such as cytokines) and reactive oxygen species from lung cells; (ii), disruption of the autonomic nervous system via interactions between inhaled particles and pulmonary nerves; and (iii) direct translocation of ultrafine particles or particle constituents into the systemic circulation.

The biological markers of low-grade inflammation include high-sensitivity C-reactive protein (hsCRP) and inflammatory cytokines. High-sensitivity C-reactive protein (hsCRP) is a marker of low-grade inflammation and is measured quite frequently in epidemiological studies; in contrast, other cytokines are less frequently assayed. CRP is one of the most distinctive acute phase reactants. Plasma levels of CRP increase rapidly and significantly in response to inflammation, cell damage, or tissue injury. Hence, CRP has long been considered as a prime marker of inflammation. High-sensitivity CRP assays can accurately measure small increases within the normal CRP range and can thus identify low but persistent levels of inflammation. Given that atherosclerosis has an inflammatory component, an hsCRP assay provides information on the cardiovascular risk (9). In studies of the underlying molecular and cellular mechanisms of inflammation, cytokine levels are generally measured in serum or plasma samples from patients or healthy individuals. Indeed, cytokines are soluble proteins that act as chemical messengers in the immune response and have an important role in cell-to-cell communication, cell growth, and cell differentiation. Moreover, abnormal secretion of specific cytokines has been associated with the pathogenesis or progression of chronic inflammatory diseases and with exposure to chemical or biological contaminants in the environment. In both large epidemiological studies and routine clinical practice, cytokine detection and quantification can usefully complement CRP assay data. In the present study, we assayed seven cytokines (interleukins (IL)-1 β , -6, -8, -10, -17A and -22, and tumor necrosis factor alpha (TNF α)) known to be involved in inflammatory diseases or the response to air pollutant exposure (10–13).

Association between air pollution and inflammation has been studied in population with a specific health condition or exposure in limited samples (from 15 to 175 subjects) and these heterogeneous studies had inconsistent results (8,14–33). Association between hsCRP and inflammation has also been studied in pregnant women showing higher hsCRP was significantly associated with mean annual exposure to particulate matter with a diameter < 10 μ m (PM₁₀) but not nitrogen dioxide (NO₂) (34). In term-born neonates, an association was observed between PM₁₀ and IL1 β (35). In asthmatic and non-asthmatic children, an association was observed between NO₂ and IL6 and IL10 (36). The literature data are nevertheless inconsistent: some researchers even observed lower levels of CRP (22,29,30) and pro-inflammatory cytokines (21,24,31,32,35) or higher levels of an anti-inflammatory cytokine (23,36) associated with higher exposure.

Only a few studies evaluated the association between inflammatory markers and air pollution in non-pregnant adults, in population-based studies; leaving the association between atmospheric pollution and inflammation in the general population insufficiently quantified. Most of these studies measured only hsCRP levels, and the results were inconsistent. A significant association between PM,

NO₂ and hsCRP was observed in older Chinese adults. However, the level of air pollution was very high at 91 µg/m³ (37).

Some studies conducted in European or American areas with very moderate pollution levels did not observe a significant association. For example, a study of 2252 people in Germany did not find significant associations between PM, nitrogen oxide (NO_x), ozone (O₃) on one hand and hsCRP on the other (29). In a study of 408 residents in Boston (USA), PM was not significantly associated with hsCRP (38). In a sample of 17000 people in the United Kingdom, an association between PM₁₀ and hsCRP was not observed (30). A study of 3860 members of the Swiss general population (39) did not find any significant association between PM₁₀ and CRP. In contrast, other studies found significant associations. Conversely, a statistically significant, positive association between 1-year pollution exposure to PM_{2.5} and PM₁₀ and the serum hsCRP level was observed in a German general population sample, with samples collected during two follow-up periods (40). Furthermore, a significant positive association between 1-year PM_{2.5} exposure and 6-month O₃ exposure and the serum hsCRP level (with repeat sampling) was found in a study of 2086 women in the USA (7). Lastly, in a cross-sectional analysis of the *Enquête Littoral Souffle Air Biologie Environnement* (ELISABET) study conducted in northern France, we previously observed an association between NO₂ on the day of blood sample collection and on the day before and the serum hsCRP level in a specific subgroup of 1506 healthy non-smokers (i.e. people without chronic or acute respiratory diseases, pulmonary medication use, airway obstruction, exposure to active or passive smoking, or inflammatory syndrome (defined as a serum hsCRP level > 10 mg/L) (41)

With regard to cytokines, fewer data have been published. In a study of 3860 individuals from the Swiss general population (a study that also included two follow-up periods), there were significant positive associations between 1-day, 1-week, 1-month, 3-month and 6-month exposure to particulate matter with an aerodynamic diameter <10 µm (PM₁₀) and inflammatory cytokines (IL-6, IL-1β, and TNFα) (39). However, hsCRP was not associated with PM₁₀ (39). These significant associations with cytokine levels have not yet been confirmed in other studies. As mentioned above, a study of 408 Boston residents (38) did not find any significant association. Furthermore, these population-based studies in adults (38,39) measured a small set of cytokines (IL-1β, IL-6, and TNFα). The objective of the present study was to explore (i) the associations between air pollutant levels and serum hsCRP levels in the ELISABET study population as a whole and (ii) the associations between air pollutant levels and plasma levels of a large set of cytokines in a subsample of non-smokers residing in the Lille urban area. We examined both short-term exposure (i.e. on the day of sample collection and on the day before) and long-term exposure (i.e. mean annual exposure at the residential address) and assayed the blood samples for biomarkers.

2. **Methods**

2.1. **Population**

The study population comprised male and female participants (aged 40-65) from the cross-sectional 2011-2013 ELISABET survey performed in northern France. All participants had lived in the same city (either Lille or Dunkirk) or its surrounding urban area for at least five years, and were recruited between January 2011 and November 2013 (42). The participants were selected from electoral rolls by random sampling, with stratification for sex, age and centre (Lille or Dunkirk), and were contacted in random order. Data were collected at home or (occasionally) during a consultation in a healthcare establishment. In all cases, a trained, registered nurse administered a detailed questionnaire. A blood sample was collected during the same visit (41).

We excluded ELISABET survey participants with a known inflammatory syndrome (defined as a serum hsCRP level > 10 mg/L) and participants with missing data for the hsCRP level, pollutant levels, or covariates.

Next, we studied other inflammatory biomarkers (IL1β, IL6, IL8, IL10, IL17A, IL22, and TNFα) in a subsample of 1125 non-smokers, in order to avoid the bias due to smoking-induced inflammation.

We included participants from the Lille area only because the standard deviation of the residential air PM₁₀ level was higher in Lille than in Dunkirk (1.936 µg/m³ versus 0.999 µg/m³, respectively), which therefore increased the statistical power of an analysis of a smaller number of participants. To enable subsequent genetic studies, we included participants of Caucasian origin with data on genetic variables. Lastly, we excluded patients with a known inflammatory syndrome, non-acceptable spirometry data, non-fasting data, or missing data.

The study was registered at ClinicalTrials.gov (NCT02490553) and the protocol was approved by the local investigational review (*CPP Nord Ouest IV*, Lille, France; reference: 2010-A00065-34), in compliance with the French legislation on biomedical research. All participants provided their written, informed consent to participation in the study (41).

2.2. Study areas

We studied participants living in two urban areas in northern France (Lille and Dunkirk, including the main city and the immediate suburbs). A map showing the cities' situation is shown in Supplemental Figure 1. Firstly, we checked that the characteristics of the two urban areas met the APHEKOM criteria, i.e. homogeneous exposure to a given air pollutant on a given day throughout the study area (43).

The Lille area is a dense urban area, with only a few non-urban zones between the three main cities (Lille, Roubaix and Tourcoing) and two smaller cities (Armentières and Comines). The road/motorway network in this area is very dense. The Lille urban area is crossed by five motorways; in 2012, the mean number of vehicles per day ranged from 36,000 to 176,000 (44,45).

The Dunkirk area is a coastal urban area. The sea is to the north of the city, and a harbour and an industrial area are located north-west of the city. The prevailing winds are south-westerly; as a consequence, air pollution from the industrial area is mostly pushed out to sea. The road and motorway network in this area is also dense (41). The Dunkirk urban area is crossed by two motorways; in 2012, the mean number of vehicles per day ranged from 27,000 to 61,000 (44,45)

2.3. Air pollution and meteorological measurements

The estimation of air pollution exposure has been described previously (41,46,47). We examined both short-term and residential exposure, using two different data sources. Short-term exposure was defined as the mean level of pollution in each study area (Lille or Dunkirk) on each measurement day and on the preceding day. This variable provides information on the short-term, reversible effects of high-pollution periods. Residential exposure was defined as the mean annual exposure at the residential address. This variable provides information on differences between people living in highly polluted areas vs. those living in less polluted areas, i.e. the long-term effects of air pollution. In brief, short-term exposure was assessed via daily measurements of PM₁₀ and nitrogen dioxide (NO₂) and hourly measurements of ozone (O₃) at air pollution monitoring stations operated by the ATMO Hauts-de-France air quality monitoring association (Lille, France). The data from the urban and suburban stations met the APHEKOM homogeneity criteria and were included. The daily pollution levels for each city area were defined as the mean daily air concentration for PM₁₀ and NO₂ and the maximal 8-hour moving average for the O₃ air concentration, as measured at monitoring stations across the urban and suburban areas. We checked that the daily measurements from monitoring stations in the same city area were correlated (41,43). Meteorological data (including daily rainfall, air temperature, relative humidity, and atmospheric pressure at sea level) were obtained from the French National Meteorological Service.

For residential exposure, the annual mean concentrations of PM₁₀ and NO₂ between 2010 and 2013 in Lille and between 2012 and 2013 in Dunkirk were taken from estimates produced by ATMO – Nord Pas de Calais, using the Atmospheric Dispersion Modelling System (ADMS, version 3.4, for urban areas) developed by Cambridge Environmental Research Consultants (48,49). This model could not be used to estimate residential exposure to O₃. The latter incorporated meteorological, topographic and land-use data, pollutant emissions from natural sources and those related to human activity, and ambient air pollution data from monitoring stations. Concentration maps for PM₁₀ and NO₂ in Lille

and Dunkirk in 2015 are shown on the ATMO website (ATMO Nord-Pas-de-Calais, 2016a, 2016b). Maps of the spatial distribution of air pollutants for 2015 in the Lille and Dunkirk urban areas are available on the ATMO Hauts-de-France website (www.atmo-hdf.fr) (50,51). We applied a spatial resolution of 25 × 25 m. Each participant's place of residence was located within a 25 m grid. We assessed the annual exposure levels at the place of residence as the mean value of the four closest points in the grid, weighted by the inverse square distance to each point (46,52).

2.4. Biomarkers

The serum concentration of [hsCRP](#) was measured in a nephelometric assay (BN [ProSpec](#) System, Siemens) with a detection range of 0.17 to 10 mg/L. Values below the limit of detection (LOD) were computed as $0.12 \text{ (LOD}/\sqrt{2}) \text{ mg/L}$ (41).

For cytokine measurements, we used the [MesoScale Discovery](#)[®] electrochemiluminescent multiplex immunoassay (Meso Scale Diagnostics LLC, Rockville, MA, USA). Multiplex assays (for IL-1 β , IL-6, IL-8, IL-10, IL-17A, and TNF α) and the individual assay (for IL22) were provided on 10-spot plates (U-PLEX Plate K15067L-2, Meso Scale Diagnostics LLC). The assays were performed according to the manufacturer's instructions, except that we modified the range of the calibration curves. The manufacturer recommends eight-point calibration curves; however, [in order to](#) improve the assay's sensitivity, we added two points within the lower part of the curve and removed two points from the upper part of the curve.

Briefly, for a single plate, 200 μL of antibodies were first incubated with 300 μL of linkers for 30 minutes. Next, 200 μL of a stop solution were added, and the mixture was incubated for 30 minutes, loaded into [QuickPlex 96-Well](#) microplates (Meso Scale Diagnostics LLC), and incubated for 1 hour at room temperature. Thereafter, 25 μL of diluent and 25 μL of samples, standards or controls were added in duplicate and incubated overnight at 4°C. Each well was washed, incubated for 1 hour with a detection antibody, washed again, and supplemented with Read Buffer (Meso Scale Diagnostics LLC). Electrochemiluminescence signals were detected with the [QuickPlex SQ 120](#) instrument and [analyzed](#) using Discovery Workbench 4.0 software (both from Meso Scale Diagnostics LLC). Values below the lower limit of quantification (the lowest concentration on the standard curve which gave a percentage coefficient of variation of less than 20%) were counted as real concentrations. The biomarker assay's stability over time was checked by running control samples on a regular basis during the study period. To this end, we pooled 40 residual plasma samples collected during routine clinical care from adult patients admitted to Lille University Hospital (Lille, France). At least one control sample was included in each run.

2.5. Statistical analysis

We looked at both short-term and residential pollution. We calculated Pearson's coefficient for the correlations with residential exposure in each year and each city. Data for residential exposure in 2011 in Dunkirk area were not available (46). We imputed the 2011 residential exposure data for this area using the 2012 residential exposure data and mean of daily measure from pollution station using the following formula. "Imputed residential exposure in 2011" = "residential exposure in 2012" – "annual mean of daily air pollution station measurement in 2012" + "annual mean of daily air pollution station measurement in 2011". [In order to](#) evaluate the accuracy of this imputation, we [imputed](#) residential exposure in 2012 with 2013 data and calculated the agreement between the imputed values and the available values using the intra-class correlation (ICC) in the whole sample of the ELISABET study.

Associations between exposure to each air pollutant and the explanatory variables (the levels of inflammatory markers) were estimated using multiple linear regression models. [In order to](#) normalize the data distribution, all explanatory variables were log-transformed. Hence, the regression coefficients were expressed as the percentage change [95% confidence interval (CI)] per 10 $\mu\text{g}/\text{m}^3$ increment in each pollutant for short-term exposure and per 2 $\mu\text{g}/\text{m}^3$ of PM_{10} or 5 $\mu\text{g}/\text{m}^3$ of NO_2 increment for residential exposure. These increments for residential exposure were chosen because there were of the same order of magnitude as the interquartile range of the measurements in Lille

and Dunkirk. The values were back-transformed and converted into a percentage change by multiplying by 100 and subtracting 100 percentage points.

We tested for interactions between pollutant levels on one hand and the urban area, smoking status, and BMI on the other.

For short-term exposure, the models were adjusted for the urban area (Lille or Dunkirk), year of inclusion, age, sex, BMI, smoking status (non-smokers/current smokers), educational level (5 or more years of higher education, 2 to 4 years of higher education, no higher education, or no secondary education), population density (calculated from the IRIS submunicipal area databases obtained from the French National Institute for Statistics and Economic Studies: www.insee.fr), meteorological measurements (humidity, atmospheric pressure, air temperature, and rainfall), season (winter, spring, summer, or autumn), day of the week, the number of days since the beginning of the study, and school vacation periods.

For residential exposure, the models were adjusted for the urban area (Lille or Dunkirk), year of inclusion, age, sex, BMI, smoking status (non-smokers/current smokers) and educational level (5 or more years of higher education, 2 to 4 years of higher education, no higher education, or no secondary education).

Our primary analysis involved tests for associations between the explanatory variables and short-term exposure to PM₁₀, NO₂ and O₃, and associations between the explanatory variables and residential exposure to PM₁₀ and NO₂. We then performed two sensitivity analyses; the first with the inclusion of PM₁₀, NO₂, and O₃ in a multipollutant model, and the second with the inclusion of wind direction and speed and cumulative hours of sunshine during the day (to determine whether the wind might bring industrial pollution into the Dunkirk area). Wind and sunshine data were obtained from the regional meteorological office (Météo France, Lille, France). Sunshine was included in the model because sunlight is involved in the formation of ground-level O₃ via the photodissociation of NO₂ (53)

The wind direction (averaged over an hour) was measured at a height of 10 meters.

We compared days with a northerly wind (between 270 and 89 degrees on the compass) with days with a southerly wind (between 90 and 269 degrees on the compass). The mean wind speed (in m/s) was calculated per hour and then per day. We excluded hourly wind speeds below 2 m/s, and calculated the hours per day with a northerly wind and the hours per day with a southerly wind. We only considered full days with a northerly wind or southerly wind for at least 70% of the day. Sunshine was considered as the cumulative duration per day. The latter variable was not measured in Dunkirk, and so we used data from the nearest station (Calais-Marck, about 30 km away). All statistical analyses were performed using R software (version 1.3.1093, R Core Team, R Foundation for Statistical Computing, <http://www.r-project.org>). The threshold for statistical significance was set to $p < 0.05$.

3. Results

3.1. Population

A total of 3275 participants were included in the ELISABET study. For the hsCRP analysis, we excluded 198 participants (119 had missing hsCRP data and 79 an inflammatory syndrome), leaving 3077. The short-term pollution analyses included 3011 participants after the exclusion of 26 participants with missing data for at least one covariate and 40 participants with missing data for pollutants (PM₁₀, NO₂ or O₃). The residential pollution analyses included 3074 participants after the exclusion of 2 participants with missing data for at least one covariate and 1 participant with missing data for pollutants (PM₁₀ or NO₂).

Of the 3275 participants, 1125 were included in the cytokine analysis. In the Lille Area, there were 1668 participants, of whom 1335 were non-smokers (either never-smokers or people having stopped smoking at least 12 months previously). Genetic data were available for 1125 of these 1335 participants. We then excluded 72 participants with non-acceptable spirometry data and 26

participants who had not fasted before blood sampling. 23 samples could not be assayed. After the exclusion of 22 participants with an inflammatory syndrome, the final sample comprised 982 participants.

The participants' characteristics are summarized in Table 1. The Lille and Dunkirk populations did not differ with regard to the main characteristics, except that the BMI was higher in the Dunkirk population and the educational level was higher in the Lille population. The participants' characteristics are described as a function of the BMI class (normal weight, overweight, or obesity) in Supplemental Table 1.

3.2. Pollution

The pollutant concentrations are summarized in Table 1. For short-term exposure, PM₁₀ and NO₂ concentrations were higher in Lille than in Dunkirk, whereas O₃ concentrations were similar in the two urban areas. For residential exposure, PM₁₀ and NO₂ concentrations were significantly higher in Lille than in Dunkirk.

We looked at the correlation between the various pollutants. For short-term exposure, PM₁₀ and NO₂ were positively correlated with each other ($r = 0.749$ in Dunkirk and 0.635 in Lille), while O₃ was negatively correlated with NO₂ ($r = -0.539$ in Dunkirk and -0.660 in Lille) and PM₁₀ ($r = -0.261$ in Dunkirk and -0.265 in Lille). For residential exposure, the coefficient for the correlation between PM₁₀ and NO₂ was 0.762 in Dunkirk and 0.840 in Lille.

For residential exposure, the Pearson coefficients for the correlation with exposure in 2012 and 2013 in Dunkirk were 0.94 for PM₁₀ and 0.97 for NO₂. For the 2012 imputed data using the 2013 pollution data, the measurement of the agreement by ICC was 0.94 for PM₁₀ and 0.89 for NO₂. The mean level of exposure in Dunkirk was $25.02 \mu\text{g}/\text{m}^3$ for PM₁₀ and $22.43 \mu\text{g}/\text{m}^3$ for NO₂ (in 2012); and $25.93 \mu\text{g}/\text{m}^3$ for PM₁₀ and $23.08 \mu\text{g}/\text{m}^3$ for NO₂ (in 2013). The correlation matrix for Lille is given in Supplemental Table 2; the correlation coefficients were very high.

3.3. Associations between pollution and inflammatory biomarkers

3.3.1. Association between pollution and hsCRP

For short-term exposure (Table 2), a $10 \mu\text{g}/\text{m}^3$ increment in PM₁₀ levels was associated with a significant increase [95%CI] in hsCRP (3.43% [$0.68; 6.25$]) in the total study population ($p=0.014$). A $10 \mu\text{g}/\text{m}^3$ increment in NO₂ levels was associated with a significant increase in hsCRP in Lille residents (7.85% [$2.04; 13.98$]; $p=0.0076$) but not in Dunkirk residents (-2.61% [$-7.55; 2.59$]; $p=0.32$; p interaction = 0.030). Furthermore, the association was significant in current smokers ($+12.37\%$ [$1.99; 23.8$]; $p=0.019$) but not in non-smokers (-0.43% [$-4.35; 3.64$]; $p=0.83$), and the interaction was statistically significant ($p=0.018$). For O₃, the association was statistically significant in Lille residents (with a 5.32% [$-8.98; -1.50$] decrease in hsCRP; $p=0.0068$) but not in Dunkirk residents ($+3.59\%$ [$-0.72; 8.09$]; $p=0.10$), and the interaction was significant ($p=0.025$).

Moreover, we observed a significant interaction between BMI and PM₁₀ ($p < 0.01$) and between BMI and NO₂ ($p=0.043$). The hsCRP level increased with PM₁₀ exposure in participants with a normal BMI and in overweight participants, although the association was only statistically significant in the latter BMI class ($p=0.0032$). No significant associations were observed in the obese class. For NO₂ exposure, the changes in hsCRP levels were similar but the associations were not significant in the normal weight and overweight BMI classes; however, a significant decrease was observed in the obese BMI class ($p=0.029$). After the exclusion of obese participants from the analysis, the associations were significant: $+6.17\%$ [$2.72; 9.73$] ($p < 0.001$) for PM₁₀ (p interaction < 0.01) and $+5.83\%$ [$1.06; 10.82$] ($p=0.016$) for NO₂ (p interaction for obese vs non-obese = 0.016). Besides, the interaction with the urban area was no more significant after adjustment for the interaction between pollution and BMI. For residential exposure (Table 2), the associations between air pollutants and the hsCRP level were not statistically significant. Likewise, the interactions between residential exposure to air pollutants and BMI were not statistically significant.

3.3.2. Association between pollution and cytokines

None of the associations between IL6, IL8, IL10, IL22, TNF, IL1, and IL17A and daily (short-term) exposure to PM₁₀ and NO₂ were significant (Table 3). For short-term exposure to O₃ only, the association with TNF α was significant (+2.69% [0.44; 4.99]; p=0.019). None of the associations with mean annual residential exposure to PM₁₀ or NO₂ were significant.

3.4. Sensitivity analyses

We performed linear regressions on multipollutant models. After adjustment for O₃, the results for short-term PM₁₀ exposure (Supplemental Table 3) followed the same trend as the primary results. After adjustment for NO₂, the PM₁₀ association was no longer significant for active smokers (+6.48% [-2.81; 16.66]; p=0.18) but became significant for non-smokers (+5.03% [0.77; 9.47]; p=0.020), with a non-significant interaction (p=0.088). For NO₂ exposure, the associations were non-significant after adjustment for PM₁₀ but the association with hsCRP was still significant in active smokers after adjustment for O₃ (+15.37% [2.97; 29.26] p=0.014), with a significant interaction (p=0.017). With regard to O₃ exposure, the associations with hsCRP levels were not significant after adjustment for PM₁₀ or NO₂. The interaction with the urban area remained significant.

For residential exposure, multipollutant models did not show any statistically significant associations with hsCRP (Supplemental Table 4).

In multipollutant models, the associations between air pollution and cytokines were not statistically significant.

We also performed a linear regression by adding meteorological adjustment covariates (wind direction, wind speed, and sunshine) to the model for short-term pollution and hsCRP levels. For PM₁₀ exposure, the associations with hsCRP were statistically significant in the total population, Lille residents, and non-smokers. The associations between NO₂ or O₃ and hsCRP were not significant. The interactions between pollutants and the urban area and between pollutants and smoking status were no longer significant. Lastly, the interactions between pollutants and the wind direction were not significant in either Lille or Dunkirk.

4. Discussion

Short-term exposure to air pollution was associated with serum hsCRP levels, and the association was influenced by smoking status, BMI, and urban area. Residential exposure was not significantly associated with serum hsCRP or plasma cytokine levels. The associations between short-term pollution and cytokine levels were not significant, with the exception of the association between O₃ and TNF α .

4.1. hsCRP

In a large survey of adults exposed to moderate levels of air pollution in two urban areas in northern France, we found that exposure on the day of the blood sample collection and the day before was significantly associated with higher hsCRP levels. The association was stronger in Lille residents, current smokers, and non-obese individuals.

Previous studies have found associations in specific, smaller populations (i.e. fewer than 100 people). For example, the blood CRP level was associated with occupational exposure among workers in a pulp and paper mill in Sweden (15) and among traffic police officers in Nepal (16). Furthermore, the CRP level in dried blood spots was associated with use of a traditional wood-burning cookstove in Nicaragua (33), and the serum CRP level was associated with exposure to metal fumes in Germany (28). A significant, positive association between daily atmospheric PM_{2.5} exposure and CRP (repeated measurements) was observed in a sample of 88 elderly residents in the US state of Utah. (14) Data for the general population are scarce. Two studies of general population samples did not find any significant associations. Green et al., (7) included 2086 women aged between 42 and 52 from six urban areas in the USA. The associations between air pollutant levels (PM_{2.5} and O₃) and the serum hsCRP level were not significant for 1-day and 30-day exposures. Likewise, a study of 3860 members of the Swiss CoLaus cohort (39) did not find significant associations between 1-day, 1-week or 1-

month PM₁₀ exposures and serum hsCRP (with two blood samples per participant). Lastly, Hennig et al.'s study (40) of 8204 people from three large adjacent cities in the Ruhr area of Germany found an association between short-term (7-day) PM_{2.5} or PM₁₀ traffic-related exposure and serum hsCRP (two blood samples per participant). However, the associations with industry-generated pollution and total pollution were not significant.

It is well established that airborne pollutants can cause lung inflammation directly through the activation of alveolar macrophages and the secretion of inflammatory cytokines, such as TNFs and ILs (54–56). The overexpression of these cytokines (predominantly IL-6) might induce an acute hepatic response, with subsequent increases in the blood levels of various proteins associated with systemic inflammation - including CRP. The blood CRP concentration is usually low under normal conditions but rises quickly and markedly in response to pro-inflammatory stimuli (57–59). The plasma half-life of CRP is less than 20 hours and does not vary according to the pathophysiological conditions. Hence, the sole determinant of the circulating CRP concentration is the rate of synthesis, which directly reflects the intensity of the stimulus. The results of several *in vitro* and *in vivo* studies have suggested that CRP levels rise in response to PM exposure (60–67).

The effects of air pollution might differ according to the duration of exposure. Indeed, short-term exposure might trigger the phagocytosis of air pollutants by alveolar macrophages, which in turn might suddenly increase inflammatory biomarker levels (54). Long-term exposure might perpetuate the inflammatory process, with consequences for chronic disease (39).

Several of the studied interactions with pollutants were statistically significant. Firstly, we observed a significant interaction between smoking status and short-term exposure to NO₂. This interaction persisted in multipollutant models. Our results suggest that the effect of pollution on hsCRP levels is greater in smokers than in non-smokers.

Pilz et al., (29) found a significant interaction between pollution and smoking. However, it is hard to draw firm conclusions because the association was stronger in former smokers than in non-smokers and current smokers. The stronger association in smokers might be due to synergistic effects of air pollution and tobacco. One hypothesis is that smokers are more susceptible to the effects of air pollution because their airways and lungs are already damaged. Indeed, smoking leads to disruption of the tight junctions and thus weakens the airways' epithelial barrier function (68). Furthermore, greater epithelial permeability in smokers might facilitate the penetration of pollutants into the body (69,70). Moreover, a decrease in pulmonary clearance in smokers might explain the greater effect of pollutants (71). Indeed, a study found that people exposed to cigarette smoke or smoking (72,73) had significantly lower plasma antioxidant levels, which might make these individuals more sensitive to the oxidative stress generated by air pollution (74). Consistently, Sun et al., (75) reported that the risk of mortality related to air pollution was greater in current or former smokers than in non-smokers.

However, a synergistic effect is not consistently observed. Panasevich et al., (76) found a significant association between air pollution and hsCRP in non-smokers but not in smokers. One hypothesis is that smokers already have a high level of low-grade inflammation due to tobacco particles, and so the additional effect of air pollution may be more difficult to detect in smokers than in non-smokers. Hence, the data on the interaction between tobacco and air pollution are sparse and inconclusive. This interaction warrants further study.

Secondly, we observed a significant interaction between pollutants (PM₁₀ and NO₂) and BMI. Thus, PM₁₀ was positively associated with hsCRP in non-obese people but not in obese people. These results are consistent with Pilz et al.'s (29) report of the association between elevated hsCRP levels and PM₁₀ in non-obese individuals but not in obese individuals. In our study (and as expected, in view of Gentile et al.'s results (77)), the hsCRP level was much higher in obese people than in people of normal weight. This was probably because obesity is associated with chronic inflammation, the activation of pro-inflammatory pathways, and higher levels of inflammatory biomarkers (78). Thus, a moderate effect of air pollution might be harder to detect in a population with a very high baseline level of low-grade inflammation.

In contrast, it has been suggested that effect of air pollution on health (not including low-grade inflammation) may be greater in obese people. For example, the association between $PM_{2.5}$ and diabetes found in the Danish Nurse cohort (79) was stronger in obese people. Surprisingly, we found NO_2 exposure was associated with significantly lower $hsCRP$ levels in obese people. We do not have any physiological hypotheses for this result, which might be due to the negative correlations between NO_2 and O_3 and the harmful effects of O_3 . Nevertheless, the association between O_3 and $hsCRP$ was not significant and did not confirm this hypothesis. This finding might also be due to chance.

Thirdly, we found interactions between short-term NO_2 and O_3 exposure and the urban area. The association between NO_2 and O_3 exposure and $hsCRP$ was statistically significant in Lille but not in Dunkirk, and the interaction was significant. The association between O_3 and $hsCRP$ was in the opposite direction to that between NO_2 and $hsCRP$. Indeed, O_3 levels and NO_2 levels are negatively correlated. There is an industrial zone in the northwest part of the Dunkirk urban area; hence, the interaction between O_3 and NO_2 levels and the urban area might be explained by a difference in the nature of the pollutants between Lille and Dunkirk. Accordingly, the association in Hennig et al.'s study (40) (where industrial and traffic-related pollutants were modelled separately) was significant only for traffic-related pollution. A northerly wind would tend to bring industrial pollution to Dunkirk, whereas a southerly wind would bring traffic-related pollution to the city and thus resemble the situation in Lille. With a southerly wind, the associations between pollution and inflammation should be similar in Lille and Dunkirk. With a northerly wind, the nature of the pollutants would differ more. Hence, one could reasonably hypothesize the presence of a pollution/wind interaction in Dunkirk. However, we did not find a significant interaction in this respect. The nature of the pollutants in Lille and Dunkirk and the differences between the two remain to be explored further. Lastly, the interaction between NO_2 and the urban area was not significant after the interaction with BMI had been taken into account – suggesting that the weaker association observed in Dunkirk (relative to Lille) was partly due to the first city's higher mean BMI.

We did not find a significant association between residential exposure and the $hsCRP$ level. This might be because most of the effects of air pollution are short-term effects. Alternatively, the difference between the highest level of PM_{10} exposure and lowest level in our moderately exposed population might be too small to be produce a significant difference. Hence, the effect of air pollution on inflammation (if present) might not differ greatly from one area of a city to another. Lastly, environmental policies tend to focus on decreasing overall pollution levels and not solely reducing the number of highly exposed subjects.

In the literature, positive associations between occupational exposure to pollution and CRP were found among 87 Pakistani automobile vehicle drivers (17), 39 taxi drivers in Brazil (20), 72 workers in petrol station in Iraq (19), 137 people exposed to diesel engine exhaust in China (21), and 78 welders in Sweden (27). Conversely, a significant negative association between occupational exposure and CRP was found in a Chinese study of 54 workers exposed to diesel engine exhaust, relative to 55 non-exposed participants (22).

Some studies have found an association between air pollution and the CRP level in particular populations, such as 175 dialysis patients in Taiwan (18), 6508 pregnant women in Netherlands (34), and 7915 older adults in China (37). Furthermore, an association was found between traffic density and the CRP level in 1017 participants of Puerto Rican origin living in Boston, USA (8).

Several studies of the general population (e.g. 3860 participants in Switzerland (39), 408 participants in Boston (USA) (38), 2252 participants in Germany (29) and about 17000 participants in the United Kingdom (30)) have not found a significant association between long-term exposure to air pollution and CRP. However, Green et al.'s (7) study of 2086 women in the USA found a significant association between $hsCRP$ and 1-year $PM_{2.5}$ exposure and 6-month O_3 exposure. Lastly, a significant association between $hsCRP$ and 1-year exposure to $PM_{2.5}$ or PM_{10} was found in 8204 German participants (40). The mean residential 365-day concentration of PM_{10} was $19.68 \mu\text{g}/\text{m}^3$ for the overall pollution, including a mean of $0.81 \mu\text{g}/\text{m}^3$ from traffic-related pollution and $2.47 \mu\text{g}/\text{m}^3$ from industrial

pollution. It is noteworthy that both of these studies featured repeated measurements of the hsCRP level.

4.2. Cytokines

The associations between plasma cytokine levels and short-term and residential exposures were not significant, with the exception of the association between short-term O₃ exposure and TNF α . Previous studies have demonstrated associations between air pollution and cytokine in highly exposed people or in specific population. For example, studies of people occupationally exposed to pollution have found positive associations (17,19,20,22) or negative associations (21,31,32) between exposure and proinflammatory cytokine levels. Studies of children with asthma (23,26,36) have found positive associations between pollution exposure and blood levels of pro-inflammatory cytokines and (surprisingly, in two reports) the anti-inflammatory cytokine IL-10 (23,36). PM₁₀ exposure was positively correlated with IL-1 β but negatively correlated with IL-6 in a cohort of newborns (35) and was positively correlated with IL6 but negatively with TNF α in pregnant women (24). In interventional studies, exposure to metal fumes was associated with higher cytokine levels (28), and the use of an air purifier was associated with lower IL-6 levels (25).

Conversely, population-based studies have published few data on moderate levels of exposure. Lane et al., (38) studied 408 Boston residents but did not find a significant association between mean annual air pollution exposure and cytokine levels. Lastly, Tsai et al., (39) studied a cohort of 3860 subjects aged from 35 to 75 years in Switzerland. Biomarker measurements were repeated twice, and PM₁₀ exposure was assigned according to the participants' home address. Short-term exposure was defined variously as exposure one day, one week and one month before the blood sampling. On average, PM₁₀ concentrations were 22.2 $\mu\text{g}/\text{m}^3$ one day before blood sampling, 24.5 $\mu\text{g}/\text{m}^3$ one week before blood sampling and 24.8 $\mu\text{g}/\text{m}^3$ one month before blood sampling. Long-term exposure was defined as exposure 3 months and 6 months before blood sampling. The concentrations of the pollutants in Tsai et al.'s study were similar to those in our study. Tsai et al., found a statistically significant association between short- and long-term exposure to PM₁₀ and levels of three cytokines (IL1 β , IL6, and TNF α); this is still the only study to have found an association with cytokine levels in the general population.

It is plausible that air pollution increases inflammation and cytokine secretion, since *in vitro* studies have evidenced the production of TNF α by alveolar macrophages exposed to particulate air pollutants (54,55). Furthermore, exposure to air pollution can lead to an increase in oxidative stress (80), a failure to eliminate apoptotic cells, and a state of chronic inflammation (81,82).

Not significant result concerning IL1 β and IL-6 seems to be inconsistent with results of Tsai et al.

With regard to the size of the effect observed by Tsai et al., (39) the power of our study for IL1 β was 99%. Furthermore, when comparing the associations for IL-6 in Tsai et al.'s study and our study, the CIs did not overlap -suggesting that there was a true difference between the two sets of results.

Indeed, the value was around 0.049 [0.029; 0.067] unit change per 1 $\mu\text{g}/\text{m}^3$ increment in PM₁₀ for the Swiss cohort vs. -0.0036 [-0.025;0.018] for our study.

In contrast to Tsai et al., we observed a positive association between O₃ and TNF α . Tsai et al., observed a positive association between TNF α and PM₁₀. In most studies, PM₁₀ and NO₂ levels are negatively correlated with O₃ levels; hence, we expected to see a positive association between TNF α and both PM₁₀ and NO₂, and a negative correlation with O₃. However, this finding should be interpreted with caution because we performed multiple statistical tests on the cytokine data.

Lastly, our mainly non-significant results for cytokines are not consistent with those of Tsai et al., despite an apparently sufficient level of statistical power. Nevertheless, the fact that we only considered a single blood sample might also explain the discrepancy between the two sets of results. Hence, further studies of the putative association between air pollution and cytokine levels are necessary.

|

4.3. Strengths and weaknesses

The present study had several strengths. Firstly, we studied a large, population-based sample. Secondly, our biomarker assays were performed with strict quality control procedures. In particular, a high-throughput analysis of plasma cytokine levels was performed on a multiplex assay platform equipped with a high-performance electrochemiluminescence detection system (QuickPlex SQ 120, Meso Scale Diagnostics LLC). As previously described (83,84), this QuickPlex instrument was superior to conventional multiplex ELISA systems in terms of sensitivity, reproducibility, throughput, and the ability to test closely related antigens simultaneously without signal reduction due to antigen competition.

Our study also had some limitations. Firstly, we did not have data on residential exposure in 2011. However, Pearson's coefficient for the correlation between residential exposure levels in 2012 and 2013 in Dunkirk was 0.94 for PM₁₀ and 0.97 for NO₂; and the ICC for the 2012 imputed data using 2013 pollution data was 0.94 for PM₁₀ and 0.89 for NO₂— suggesting that the geographical distribution of mean annual air pollution levels is stable over time and that the imputation of residential exposure data in Dunkirk may only induce a limited classification bias. In addition, adjustment of the analysis for the year of inclusion might have reduced this bias. Furthermore, as defined in the present study, residential exposure depended mainly on the household's geographical location and less on the measurement timepoint. Therefore, the bias due to the use of data on geographical disparities from 2012 (rather than 2011) might be limited. Secondly, the difference in residential exposure between the most and least exposed participants was small. For example, the interquartile range was 25.84-27.19 µg/m³ in Lille; this might have limited the study's power for the analysis of residential exposure. Thirdly, a lack of precision in the measurement of pollutant concentrations might have induced non-differential bias and reduces the study's power. Fourthly, our study's cross-sectional design (with a single blood sample) prevented us from controlling for the intra-individual variability in cytokine levels. In comparison, studies with repeated biomarker measurements have shown significant associations between pollution exposure and cytokine (39) or hsCRP (40) levels. Nevertheless, our statistically significant results were in line with known associations (85–87), such as the associations with BMI, IL-6 ($p < 10^{-17}$), TNF ($p < 10^{-3}$) and the association between pulmonary obstruction and IL-6 ($p < 0.05$) (data not shown). This confirmed our ability to evidence a strong association, even when cytokine levels varied significantly. Lastly, our population sample size might not have been large enough for some analyses (e.g. the cytokine analysis in particular). For example, the association between pollution and hsCRP in the subsample of 982 non-smokers from Lille was not significant (data not shown) - suggesting the need for a larger sample.

5. Conclusion

We observed a significant association between short-term exposure to air pollution and the serum hsCRP level; this association is consistent with the hypothesis whereby the health effects of air pollution are mediated by low-grade inflammation. This finding is consistent with some (but not all) previous observations. We found that BMI and smoking interacted with the association between short-term exposure to air pollution and the serum hsCRP level. This interaction might explain discrepancies in the literature data and might need to be taken into account in studies of hsCRP and pollution. We studied a large set of cytokines; all but one of the associations between air pollution exposure and cytokine levels were non-significant and so were not in agreement with Tsai et al.'s previous report (39). Hence, further research on this topic must include precise air pollutant measurements and must take account of intra-individual variability in cytokine levels.

Acknowledgements

We thank Florent Occelli for the map of northern France. We thank the ATMO Hauts-de-France air quality monitoring association for air pollution measurements and modelling in the Lille and Dunkirk

urban areas. We also thank Lille University Hospital (especially the Centre de Biologie et de Pathologie), the University of Lille, the Institut Pasteur of Lille (especially the Departments of Médecine Préventive, Biologie Spécialisée and Médecine du Travail, and the Laboratoire d'Analyses Génomiques) and the Centre Hospitalier Général de Dunkerque (especially the Departments of Biology and Pneumology). We particularly thank the nurses, physicians and secretarial staff at the University of Lille and the Institut Pasteur of Lille. Lastly, we thank the French Ministère de l'Enseignement Supérieur et de la Recherche, the Hauts de France Region and the European Regional Development Fund for financial support.

Funding

This work was funded by Lille University Hospital (CHU de Lille, Lille, France), the Nord Pas-de-Calais Regional Council, and the European Regional Development Fund (ERDF-FEDER Presage N°36034) as part of the CPER Institut de Recherche en Environnement Industriel (IRENI) program. The work was also supported by the French government through the Programme Investissement d'Avenir (I-SITE ULNE / ANR-16-IDEX-0004 ULNE), managed by the Agence Nationale de la Recherche as part of the project "Santé Environnement: du risque territorial au risque individuel" (also funded by Lille European Metropole). Lastly, this work is part of the CPER's CLIMIBIO research project.

Competing interest

LD, RM and JMLG have contributed to an expert report (commissioned by Lille European Metropole) entitled "Rapport d'expertise à propos de la localisation de la piscine du projet d'aménagement de la gare Saint Sauveur à Lille" [Expert report on the location of the swimming pool in the Saint Sauveur station development project in Lille] but did not receive any personal fees.

Other authors declare that they have no conflicts of interest.

References

1. Air pollution [Internet]. [cited 2021 May 19]. Available from: <https://www.who.int/news-room/air-pollution>
2. Brook RD, Rajagopalan S, Pope CA, Brook JR, Bhatnagar A, Diez-Roux AV, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*. 2010 Jun 1;121(21):2331–78.
3. Pearson TA, Mensah GA, Alexander RW, Anderson JL, Cannon RO, Criqui M, et al. Markers of inflammation and cardiovascular disease: application to clinical and public health practice: A statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. *Circulation*. 2003 Jan 28;107(3):499–511.
4. León-Pedroza JI, González-Tapia LA, del Olmo-Gil E, Castellanos-Rodríguez D, Escobedo G, González-Chávez A. Inflamación sistémica de grado bajo y su relación con el desarrollo de enfermedades metabólicas: de la evidencia molecular a la aplicación clínica. *Cir Cir*. 2015 Nov 1;83(6):543–51.
5. Lee HM, Cha JM, Lee JL, Jeon JW, Shin HP, Joo KR, et al. High C-reactive protein level is associated with high-risk adenoma. *Intest Res*. 2017 Oct;15(4):511–7.
6. Leonard BE. Inflammation and depression: a causal or coincidental link to the pathophysiology? *Acta Neuropsychiatr*. 2018 Feb;30(1):1–16.

7. Green R, Broadwin R, Malig B, Basu R, Gold EB, Qi L, et al. Long- and Short-term Exposure to Air Pollution and Inflammatory/Hemostatic Markers in Midlife Women. *Epidemiol Camb Mass*. 2016 Mar;27(2):211–20.
8. Rioux CL, Tucker KL, Mwamburi M, Gute DM, Cohen SA, Brugge D. Residential traffic exposure, pulse pressure, and C-reactive protein: consistency and contrast among exposure characterization methods. *Environ Health Perspect*. 2010 Jun;118(6):803–11.
9. Rifai N, Ridker PM. High-sensitivity C-reactive protein: a novel and promising marker of coronary heart disease. *Clin Chem*. 2001 Mar;47(3):403–11.
10. Barnes PJ. Cellular and molecular mechanisms of asthma and COPD. *Clin Sci Lond Engl* 1979. 2017 Jul 1;131(13):1541–58.
11. Akiki Z, Rava M, Diaz Gil O, Pin I, le Moual N, Siroux V, et al. Serum cytokine profiles as predictors of asthma control in adults from the EGEA study. *Respir Med*. 2017 Apr;125:57–64.
12. Shamsollahi HR, Jahanbin B, Rafieian S, Yunesian M. Particulates induced lung inflammation and its consequences in the development of restrictive and obstructive lung diseases: a systematic review. *Environ Sci Pollut Res Int*. 2021 May;28(20):25035–50.
13. Plé C, Fan Y, Ait Yahia S, Vorng H, Everaere L, Chenivresse C, et al. Polycyclic Aromatic Hydrocarbons Reciprocally Regulate IL-22 and IL-17 Cytokines in Peripheral Blood Mononuclear Cells from Both Healthy and Asthmatic Subjects. *PLoS ONE*. 2015 Apr 10;10(4):e0122372.
14. Pope CA, Hansen ML, Long RW, Nielsen KR, Eatough NL, Wilson WE, et al. Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environ Health Perspect*. 2004 Mar;112(3):339–45.
15. Westberg H, Elihn K, Andersson E, Persson B, Andersson L, Bryngelsson I-L, et al. Inflammatory markers and exposure to airborne particles among workers in a Swedish pulp and paper mill. *Int Arch Occup Environ Health*. 2016 Jul;89(5):813–22.
16. Shakya KM, Peltier RE, Zhang Y, Pandey BD. Roadside Exposure and Inflammation Biomarkers among a Cohort of Traffic Police in Kathmandu, Nepal. *Int J Environ Res Public Health*. 2019 Jan 29;16(3).
17. Riaz H, Syed BM, Laghari Z, Pirzada S. Analysis of inflammatory markers in apparently healthy automobile vehicle drivers in response to exposure to traffic pollution fumes. *Pak J Med Sci*. 2020 Jun;36(4):657–62.
18. Huang W-H, Yen T-H, Chan M-J, Su Y-J. Environmental carbon monoxide level is associated with the level of high-sensitivity C-reactive protein in peritoneal dialysis patients. *Medicine (Baltimore)*. 2014 Nov;93(26):e181.
19. Sajid Jabbar A, Ali ET. Impact of Petroleum Exposure on Some Hematological Indices, Interleukin-6, and Inflammatory Markers of Workers at Petroleum Stations in Basra City. *J Environ Public Health*. 2020;2020:7693891.
20. Brucker N, Moro AM, Charão MF, Durgante J, Freitas F, Baierle M, et al. Biomarkers of occupational exposure to air pollution, inflammation and oxidative damage in taxi drivers. *Sci Total Environ*. 2013 Oct 1;463–464:884–93.

21. Dai Y, Zhang X, Zhang R, Zhao X, Duan H, Niu Y, et al. Long-term exposure to diesel engine exhaust affects cytokine expression among occupational population. *Toxicol Res.* 2016 Mar 1;5(2):674–81.
22. Bassig BA, Dai Y, Vermeulen R, Ren D, Hu W, Duan H, et al. Occupational exposure to diesel engine exhaust and alterations in immune/inflammatory markers: a cross-sectional molecular epidemiology study in China. *Carcinogenesis.* 2017 Oct 26;38(11):1104–11.
23. Negherbon JP, Romero K, Williams DL, Guerrero-Preston RE, Hartung T, Scott AL, et al. Whole Blood Cytokine Response to Local Traffic-Related Particulate Matter in Peruvian Children With and Without Asthma. *Front Pharmacol.* 2017;8:157.
24. Buxton MA, Meraz-Cruz N, Sanchez BN, Gronlund CJ, Foxman B, Vadillo-Ortega F, et al. Air pollution and inflammation: Findings from concurrent repeated measures of systemic and reproductive tract cytokines during term pregnancy in Mexico City. *Sci Total Environ.* 2019 Sep 1;681:235–41.
25. Chen R, Li H, Cai J, Wang C, Lin Z, Liu C, et al. Fine Particulate Air Pollution and the Expression of microRNAs and Circulating Cytokines Relevant to Inflammation, Coagulation, and Vasoconstriction. *Environ Health Perspect.* 2018 Jan 17;126(1):017007.
26. Klümper C, Krämer U, Lehmann I, von Berg A, Berdel D, Herberth G, et al. Air pollution and cytokine responsiveness in asthmatic and non-asthmatic children. *Environ Res.* 2015 Apr;138:381–90.
27. Taj T, Gliga AR, Hedmer M, Wahlberg K, Assarsson E, Lundh T, et al. Effect of welding fumes on the cardiovascular system: a six-year longitudinal study. *Scand J Work Environ Health.* 2021 Jan 1;47(1):52–61.
28. Baumann R, Joraslafsky S, Markert A, Rack I, Davatgarbenam S, Kossack V, et al. IL-6, a central acute-phase mediator, as an early biomarker for exposure to zinc-based metal fumes. *Toxicology.* 2016 Dec 12;373:63–73.
29. Pilz V, Wolf K, Breitner S, Ruckerl R, Koenig W, Rathmann W, et al. C-reactive protein (CRP) and long-term air pollution with a focus on ultrafine particles. *Int J Hyg Environ Health.* 2018 Apr;221(3):510–8.
30. Forbes LJJ, Patel MD, Rudnicka AR, Cook DG, Bush T, Stedman JR, et al. Chronic exposure to outdoor air pollution and markers of systemic inflammation. *Epidemiol Camb Mass.* 2009 Mar;20(2):245–53.
31. Matsuda M, Bonatti R, Marquezini MV, Garcia MLB, Santos UP, Braga ALF, et al. Lacrimal Cytokines Assessment in Subjects Exposed to Different Levels of Ambient Air Pollution in a Large Metropolitan Area. *PLoS One.* 2015;10(11):e0143131.
32. Dai Y, Ren D, Bassig BA, Vermeulen R, Hu W, Niu Y, et al. Occupational exposure to diesel engine exhaust and serum cytokine levels. *Environ Mol Mutagen.* 2018 Mar;59(2):144–50.
33. Young BN, Peel JL, Nelson TL, Bachand AM, Heiderscheidt JM, Luna B, et al. C-reactive protein from dried blood spots: Application to household air pollution field studies. *Indoor Air.* 2020 Jan;30(1):24–30.

34. van den Hooven EH, de Kluizenaar Y, Pierik FH, Hofman A, van Ratingen SW, Zandveld PYJ, et al. Chronic air pollution exposure during pregnancy and maternal and fetal C-reactive protein levels: the Generation R Study. *Environ Health Perspect*. 2012 May;120(5):746–51.
35. Latzin P, Frey U, Armann J, Kieninger E, Fuchs O, Rössli M, et al. Exposure to moderate air pollution during late pregnancy and cord blood cytokine secretion in healthy neonates. *PLoS One*. 2011;6(8):e23130.
36. Gruziova O, Merid SK, Gref A, Gajulapuri A, Lemonnier N, Ballereau S, et al. Exposure to Traffic-Related Air Pollution and Serum Inflammatory Cytokines in Children. *Environ Health Perspect*. 2017 Jun 16;125(6):067007.
37. Elbarbary M, Oganessian A, Honda T, Morgan G, Guo Y, Guo Y, et al. Systemic Inflammation (C-Reactive Protein) in Older Chinese Adults Is Associated with Long-Term Exposure to Ambient Air Pollution. *Int J Environ Res Public Health*. 2021 Mar 22;18(6):3258.
38. Lane KJ, Levy JJ, Scammell MK, Peters JL, Patton AP, Reisner E, et al. Association of modeled long-term personal exposure to ultrafine particles with inflammatory and coagulation biomarkers. *Environ Int*. 2016 Aug;92–93:173–82.
39. Tsai D-H, Riediker M, Berchet A, Paccaud F, Waeber G, Vollenweider P, et al. Effects of short- and long-term exposures to particulate matter on inflammatory marker levels in the general population. *Environ Sci Pollut Res Int*. 2019 Jul;26(19):19697–704.
40. Hennig F, Fuks K, Moebus S, Weinmayr G, Memmesheimer M, Jakobs H, et al. Association between source-specific particulate matter air pollution and hs-CRP: local traffic and industrial emissions. *Environ Health Perspect*. 2014 Jul;122(7):703–10.
41. Dauchet L, Hulo S, Cherot-Kornobis N, Matran R, Amouyel P, Edmé J-L, et al. Short-term exposure to air pollution: Associations with lung function and inflammatory markers in non-smoking, healthy adults. *Environ Int*. 2018 Dec;121:610–9.
42. Quach A, Giovannelli J, Chérot-Kornobis N, Ciuchete A, Clément G, Matran R, et al. Prevalence and underdiagnosis of airway obstruction among middle-aged adults in northern France: The ELISABET study 2011-2013. *Respir Med*. 2015 Dec;109(12):1553–61.
43. Pascal M, Corso M, Ung A, Declercq C, Medina S. Guidelines for assessing the health impacts of air pollution in European cities Work Package 5 Deliverable D5. :40.
44. Cartes de trafics annuels - DREAL HAUTS-DE-FRANCE [Internet]. [cited 2021 Dec 28]. Available from: <https://www.hauts-de-france.developpement-durable.gouv.fr/?Cartes-de-trafics-annuels>
45. https://www.hauts-de-france.developpement-durable.gouv.fr/IMG/pdf/recensement_circulation_region_nord_pas_de_calais_2012_trafic_to_us_vehicules.pdf.
46. Riant M, Meirhaeghe A, Giovannelli J, Occelli F, Havet A, Cuny D, et al. Associations between long-term exposure to air pollution, glycosylated hemoglobin, fasting blood glucose and diabetes mellitus in northern France. *Environ Int*. 2018 Nov;120:121–9.
47. Havet A, Hulo S, Cuny D, Riant M, Occelli F, Cherot-Kornobis N, et al. Residential exposure to outdoor air pollution and adult lung function, with focus on small airway obstruction. *Environ Res*. 2020 Apr;183:109161.

48. CERC > Environmental software > ADMS-Urban model [Internet]. [cited 2021 Dec 28]. Available from: <http://www.cerc.co.uk/environmental-software/ADMS-Urban-model.html>
49. NUMTECH - Logiciels gamme ADMS - Conseils, vente exclusive, formation et assistance [Internet]. [cited 2021 Dec 28]. Available from: <http://www.numtech.fr/log.php?rub=urban>
50. ATMO Nord-Pas-de-Calais. Bilan territorial 2015 - Métropole Européenne de Lille [Internet]. 2016 [cited 2018 Jul 11]. Available from: http://www.atmo-hdf.fr/joomlatools-files/docman-files/Bilans-territoriaux/BT2015_MEL_VF.pdf.
51. ATMO Nord-Pas-de-Calais. Bilan territorial 2015 - Communauté Urbaine de Dunkerque [Internet]. 2016 [cited 2018 Jul 11]. Available from: http://www.atmo-hdf.fr/joomlatools-files/docman-files/BT2015_CUD_VF.compressed.pdf.
52. Babak O, Deutsch CV. Statistical approach to inverse distance interpolation. *Stoch Environ Res Risk Assess*. 2009 Jul 1;23(5):543–53.
53. Xiong Y, Du K. Source-resolved attribution of ground-level ozone formation potential from VOC emissions in Metropolitan Vancouver, BC. *Sci Total Environ*. 2020 Jun 15;721:137698.
54. van Eeden SF, Tan WC, Suwa T, Mukae H, Terashima T, Fujii T, et al. Cytokines involved in the systemic inflammatory response induced by exposure to particulate matter air pollutants (PM₁₀). *Am J Respir Crit Care Med*. 2001 Sep 1;164(5):826–30.
55. Becker S, Mundandhara S, Devlin RB, Madden M. Regulation of cytokine production in human alveolar macrophages and airway epithelial cells in response to ambient air pollution particles: further mechanistic studies. *Toxicol Appl Pharmacol*. 2005 Sep 1;207(2 Suppl):269–75.
56. Glencross DA, Ho T-R, Camiña N, Hawrylowicz CM, Pfeffer PE. Air pollution and its effects on the immune system. *Free Radic Biol Med*. 2020 May;151:56–68.
57. Gabay C, Kushner I. Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med*. 1999 Feb 11;340(6):448–54.
58. Szalai AJ, Agrawal A, Greenhough TJ, Volanakis JE. C-reactive protein: structural biology and host defense function. *Clin Chem Lab Med*. 1999 Mar;37(3):265–70.
59. Sheriff A, Kayser S, Brunner P, Vogt B. C-Reactive Protein Triggers Cell Death in Ischemic Cells. *Front Immunol*. 2021 Feb 10;12:630430.
60. Watterson TL, Hamilton B, Martin R, Coulombe RA. Urban particulate matter causes ER stress and the unfolded protein response in human lung cells. *Toxicol Sci Off J Soc Toxicol*. 2009 Nov;112(1):111–22.
61. Ramage L, Guy K. Expression of C-reactive protein and heat-shock protein-70 in the lung epithelial cell line A549, in response to PM₁₀ exposure. *Inhal Toxicol*. 2004 Jun;16(6–7):447–52.
62. Ramage L, Proudfoot L, Guy K. Expression of C-reactive protein in human lung epithelial cells and upregulation by cytokines and carbon particles. *Inhal Toxicol*. 2004 Aug;16(9):607–13.
63. Vogel CFA, Sciuillo E, Wong P, Kuzmicky P, Kado N, Matsumura F. Induction of proinflammatory cytokines and C-reactive protein in human macrophage cell line U937 exposed to air pollution particulates. *Environ Health Perspect*. 2005 Nov;113(11):1536–41.

64. Niwa Y, Hiura Y, Sawamura H, Iwai N. Inhalation exposure to carbon black induces inflammatory response in rats. *Circ J Off J Jpn Circ Soc.* 2008 Jan;72(1):144–9.
65. Rohr AC, Wagner JG, Morishita M, Kamal A, Keeler GJ, Harkema JR. Cardiopulmonary responses in spontaneously hypertensive and Wistar-Kyoto rats exposed to concentrated ambient particles from Detroit, Michigan. *Inhal Toxicol.* 2010 May;22(6):522–33.
66. Upadhyay S, Ganguly K, Stoeger T, Semmler-Bhenke M, Takenaka S, Kreyling WG, et al. Cardiovascular and inflammatory effects of intratracheally instilled ambient dust from Augsburg, Germany, in spontaneously hypertensive rats (SHRs). *Part Fibre Toxicol.* 2010 Sep 29;7:27.
67. Lei Y-C, Hwang J-S, Chan C-C, Lee C-T, Cheng T-J. Enhanced oxidative stress and endothelial dysfunction in streptozotocin-diabetic rats exposed to fine particles. *Environ Res.* 2005 Nov;99(3):335–43.
68. Heijink IH, Brandenburg SM, Postma DS, van Oosterhout AJM. Cigarette smoke impairs airway epithelial barrier function and cell-cell contact recovery. *Eur Respir J.* 2012 Feb;39(2):419–28.
69. Jones JG, Minty BD, Lawler P, Hulands G, Crawley JC, Veall N. Increased alveolar epithelial permeability in cigarette smokers. *Lancet Lond Engl.* 1980 Jan 12;1(8159):66–8.
70. Danov O, Wolff M, Bartel S, Böhlen S, Obernolte H, Wronski S, et al. Cigarette Smoke Affects Dendritic Cell Populations, Epithelial Barrier Function, and the Immune Response to Viral Infection With H1N1. *Front Med.* 2020;7:571003.
71. Mehta H, Nazzari K, Sadikot RT. Cigarette smoking and innate immunity. *Inflamm Res Off J Eur Histamine Res Soc Al.* 2008 Nov;57(11):497–503.
72. Dietrich M, Block G, Norkus EP, Hudes M, Traber MG, Cross CE, et al. Smoking and exposure to environmental tobacco smoke decrease some plasma antioxidants and increase gamma-tocopherol in vivo after adjustment for dietary antioxidant intakes. *Am J Clin Nutr.* 2003 Jan;77(1):160–6.
73. Hemilä H. The effect of β -carotene on the mortality of male smokers is modified by smoking and by vitamins C and E: evidence against a uniform effect of nutrient. *J Nutr Sci.* 2020 Mar 11;9:e11.
74. Münzel T, Sørensen M, Gori T, Schmidt FP, Rao X, Brook FR, et al. Environmental stressors and cardio-metabolic disease: part II-mechanistic insights. *Eur Heart J.* 2017 Feb 21;38(8):557–64.
75. Sun S, Cao W, Chan K-P, Ran J, Ge Y, Zhang Y, et al. Cigarette smoking increases deaths associated with air pollution in Hong Kong. *Atmos Environ.* 2020 Feb 15;223:117266.
76. Panasevich S, Leander K, Rosenlund M, Ljungman P, Bellander T, de Faire U, et al. Associations of long- and short-term air pollution exposure with markers of inflammation and coagulation in a population sample. *Occup Environ Med.* 2009 Nov;66(11):747–53.
77. Gentile M, Panico S, Rubba F, Mattiello A, Chiodini P, Jossa F, et al. Obesity, overweight, and weight gain over adult life are main determinants of elevated hs-CRP in a cohort of Mediterranean women. *Eur J Clin Nutr.* 2010 Aug;64(8):873–8.

78. Bastard J-P, Maachi M, Lagathu C, Kim MJ, Caron M, Vidal H, et al. Recent advances in the relationship between obesity, inflammation, and insulin resistance. *Eur Cytokine Netw*. 2006 Mar;17(1):4–12.
79. Hansen AB, Ravnskjaer L, Loft S, Andersen KK, Brauner EV, Baastrup R, et al. Long-term exposure to fine particulate matter and incidence of diabetes in the Danish Nurse Cohort. *Environ Int*. 2016 May;91:243–50.
80. Imrich A, Ning Y, Lawrence J, Coull B, Gitin E, Knutson M, et al. Alveolar macrophage cytokine response to air pollution particles: oxidant mechanisms. *Toxicol Appl Pharmacol*. 2007 Feb 1;218(3):256–64.
81. Grabiec AM, Hussell T. The role of airway macrophages in apoptotic cell clearance following acute and chronic lung inflammation. *Semin Immunopathol*. 2016 Jul;38(4):409–23.
82. Robb CT, Regan KH, Dorward DA, Rossi AG. Key mechanisms governing resolution of lung inflammation. *Semin Immunopathol*. 2016 Jul;38(4):425–48.
83. Bolton J, Chaudhury S, MacGill RS, Early AM, King CR, Locke E, et al. Multiplex serological assay for establishing serological profiles of polymorphic, closely related peptide antigens. *MethodsX*. 2021;8:101345.
84. Bolton JS, Chaudhury S, Dutta S, Gregory S, Locke E, Pierson T, et al. Comparison of ELISA with electro-chemiluminescence technology for the qualitative and quantitative assessment of serological responses to vaccination. *Malar J*. 2020 Apr 17;19(1):159.
85. Celli BR, Locantore N, Yates J, Tal-Singer R, Miller BE, Bakke P, et al. Inflammatory biomarkers improve clinical prediction of mortality in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012 May 15;185(10):1065–72.
86. Huang H, Huang X, Zeng K, Deng F, Lin C, Huang W. Interleukin-6 is a Strong Predictor of the Frequency of COPD Exacerbation Within 1 Year. *Int J Chron Obstruct Pulmon Dis*. 2021;16:2945–51.
87. De Lorenzo A, Del Gobbo V, Premrov MG, Bigioni M, Galvano F, Di Renzo L. Normal-weight obese syndrome: early inflammation? *Am J Clin Nutr*. 2007 Jan;85(1):40–5.

Table 1. Characteristics of the study participants

	hsCRP population N = 3074		Cytokine population N = 982
	Dunkirk	Lille	
n (%)	1508 (49.06 %)	1566 (50.94 %)	982 (100 %)
Age (years), mean (SD)	53.26 (7.26)	53.16 (7.20)	53.84 (7.17)
Sex (H), n (%)	735 (48.74 %)	742 (47.38 %)	451 (45.93 %)
Non-smokers, n (%)	1229 (81.50 %)	1279 (81.67 %)	1004 (100 %)
BMI (kg/m ²), median (SD)	27.38 (5.08)	26.54 (4.96)	26.72 (5.03)
Educational level, n (%)			
5 or more years of higher education	193 (12.80 %)	376 (24.01 %)	263 (26.78 %)
2 to 4 years of higher education	222 (14.72 %)	340 (21.71 %)	222 (22.61 %)
no higher education	902 (59.81 %)	716 (45.72 %)	422 (42.97 %)
no secondary education	191 (12.67 %)	134 (8.56 %)	75 (7.64 %)
hsCRP (mg/L), median [IQR]	1.14 [0.58-2.47]	0.99 [0.50-2.04]	0.94 [0.49-1.89]
Biomarkers (pg/mL), median [IQR]			
IL1 β			0.13 [0.13-0.13]
IL6			0.44 [0.30-0.67]
IL8			2.72 [2.09-3.76]
IL10			0.19 [0.13-0.28]
IL17a			1.39 [1.39-1.39]
TNF α			2.35 [1.92-2.90]
IL22			0.23 [0.13-0.42]
Pollutants			
Short-term PM ₁₀ ($\mu\text{g}/\text{m}^3$), median [IQR]	21.50 [17.23-33.29]	22.50 [16.5-35.82]	22.56 [16.5-35.82]
Short-term NO ₂ ($\mu\text{g}/\text{m}^3$), median [IQR]	20.25 [13.75-29.5]	24.00 [18-32.72]	24.00 [18.12-32.44]
Short-term O ₃ ($\mu\text{g}/\text{m}^3$), median [IQR]	44.96 [32.84-55.06]	44.50 [30.97-56.51]	44.90 [31.16-56.43]
Long-term PM ₁₀ ($\mu\text{g}/\text{m}^3$), median [IQR]	26.54 [25.61-27.53]	27.03 [25.7-28.28]	26.87 [25.44-28.13]
Long-term NO ₂ ($\mu\text{g}/\text{m}^3$), median [IQR]	22.12 [19.68-26.10]	26.01 [22.62-28.63]	25.53 [22.15-28.28]

Abbreviations: SD = standard deviation; BMI = body mass index; hsCRP = high sensitivity C-reactive protein; PM₁₀ = particulate matter with an aerodynamic diameter below 10 μm ; NO₂ = nitrogen dioxide; O₃ = ozone

Table 2. Associations between exposure to atmospheric pollutants and serum hsCRP (mg/L)

Population	Number	PM ₁₀			NO ₂			O ₃		
		short-term exposure			short-term exposure			short-term exposure		
		Percentage change for 10 $\mu\text{g}/\text{m}^3$ [95%CI]	p trend	p interaction	Percentage change for 10 $\mu\text{g}/\text{m}^3$ [95%CI]	p trend	p interaction	Percentage change for 10 $\mu\text{g}/\text{m}^3$ [95%CI]	p trend	p interaction
Total	3011	3.43 [0.68 ; 6.25]	0.014		1.75 [-1.96 ; 5.61]	0.36		-1.20 [-3.95 ; 1.64]	0.40	
City				0.074			0.030			0.025
Lille	1554	5.83 [2.26 ; 9.54]	0.0013		7.85 [2.04 ; 13.98]	0.0076		-5.32 [-8.98 ; -1.50]	0.0088	
Dunkirk	1457	0.16 [-4.21 ; 4.73]	0.94		-2.61 [-7.55 ; 2.59]	0.32		3.99 [-0.72 ; 8.09]	0.10	
Tobacco				0.082			0.018			0.84
Non-smokers	2461	2.22 [-0.71 ; 5.24]	0.14		-0.43 [-4.35 ; 3.64]	0.83		-1.01 [-3.99 ; 2.05]	0.51	
Current smokers	550	9.41 [2.16 ; 17.17]	0.010		12.37 [1.99 ; 23.80]	0.019		-1.82 [-8.78 ; 5.68]	0.63	
BMI				0.0074			0.043			0.57
BMI < 25	1163	3.15 [-1.74 ; 8.29]	0.21		3.34 [-3.38 ; 10.52]	0.34		-2.74 [-7.57 ; 2.34]	0.28	
25 \leq BMI < 30	1167	6.69 [2.20 ; 11.37]	0.0032		4.26 [-1.86 ; 10.76]	0.18		-0.62 [-5.05 ; 4.01]	0.79	
BMI \geq 30	681	-4.18 [-9.02 ; 0.92]	0.11		-7.47 [-13.66 ; -0.83]	0.029		2.79 [-2.59 ; 8.47]	0.32	
residential exposure										
Population	Number	Percentage change for 2 $\mu\text{g}/\text{m}^3$ [95%CI]	p trend	p interaction	Percentage change for 5 $\mu\text{g}/\text{m}^3$ [95%CI]	p trend	p interaction			
Total	3074	1.98 [-2.35 ; 6.51]	0.38		0.88 [-3.19 ; 5.13]	0.68				
City				0.27			0.12			
Lille	1566	0.13 [-5.00 ; 5.55]	0.96		-0.91 [-5.27 ; 3.66]	0.69				
Dunkirk	1508	7.27 [-3.17 ; 18.83]	0.18		9.50 [-1.66 ; 21.92]	0.098				
Tobacco				0.18			0.082			
Non-smokers	2508	3.24 [-1.42 ; 8.13]	0.18		2.84 [-1.56 ; 7.43]	0.21				
Current smokers	566	-3.44 [-14.46 ; 9.00]	0.57		-8.92 [-18.98 ; 2.38]	0.12				
BMI				0.28			0.18			
BMI < 25	1187	-0.23 [-7.31 ; 7.40]	0.95		-1.25 [-7.69 ; 5.65]	0.72				
25 \leq BMI < 30	1190	2.93 [-3.83 ; 10.17]	0.40		3.76 [-3.03 ; 11.03]	0.29				
BMI \geq 30	697	5.42 [-3.14 ; 14.72]	0.22		2.38 [-5.40 ; 10.8]	0.56				

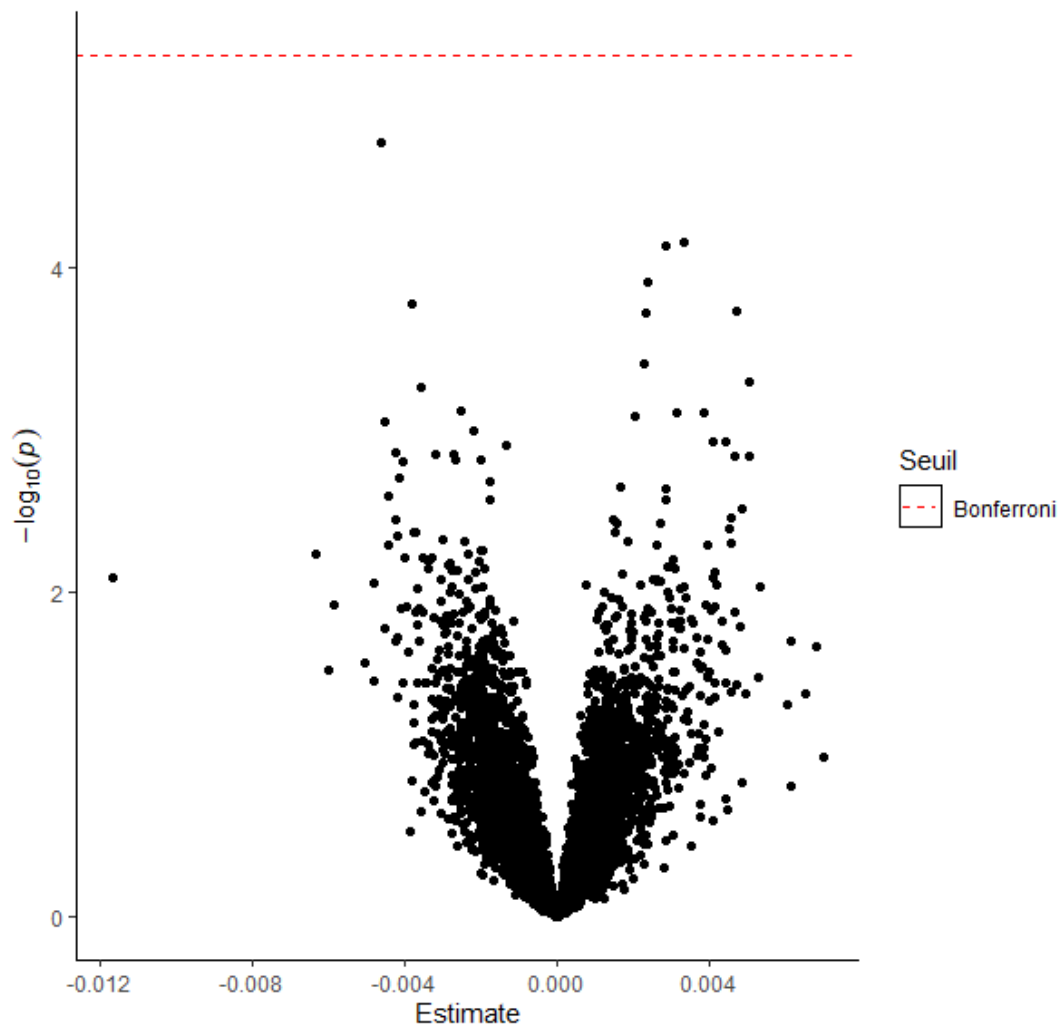
Abbreviations: PM₁₀ = particulate matter with an aerodynamic diameter below 10 μm ; NO₂ = nitrogen dioxide; O₃ = ozone; BMI = body mass index (kg/m²)

Table 3. Associations between air pollution exposure and plasma biomarker levels (pg/mL)

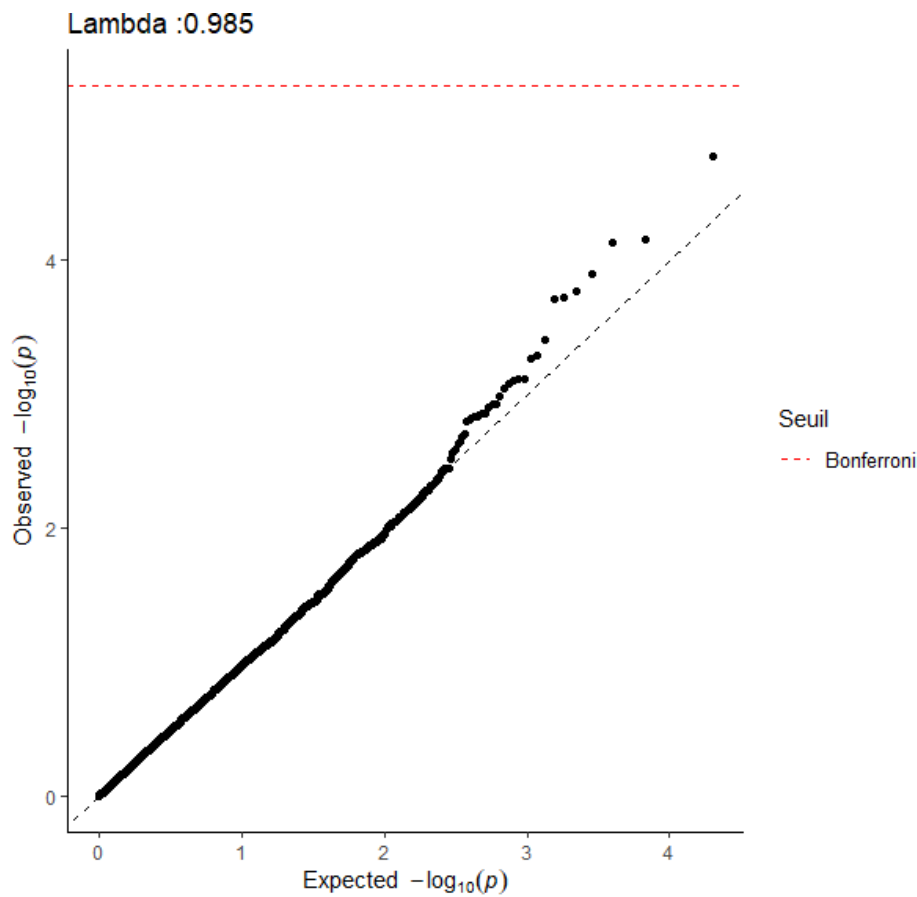
Biomarker	PM ₁₀		NO ₂		O ₃	
	<i>short-term exposure</i>					
	Percentage change for 10 µg/m ³ [95%CI]	p trend	Percentage change for 10 µg/m ³ [95%CI]	p trend	Percentage change for 10 µg/m ³ [95%CI]	p trend
IL6	1.56 [-1.44 ; 4.65]	0.31	0.67 [-4.00 ; 5.57]	0.78	1.04 [-2.31 ; 4.52]	0.55
IL8	-1.39 [-3.64 ; 0.91]	0.23	-1.84 [-5.37 ; 1.83]	0.32	1.05 [-1.56 ; 3.72]	0.43
IL10	1.87 [-1.32 ; 5.16]	0.25	0.27 [-4.67 ; 5.47]	0.92	0.16 [-3.39 ; 3.83]	0.93
IL22	1.43 [-4.06 ; 7.24]	0.62	1.76 [-6.85 ; 11.16]	0.70	2.61 [-3.65 ; 9.27]	0.42
TNF	-0.72 [-2.65 ; 1.25]	0.47	-2.25 [-5.25 ; 0.83]	0.15	2.69 [0.44 ; 4.99]	0.019
IL1	-1.81 [-6.68 ; 3.32]	0.48	-1.25 [-8.96 ; 7.10]	0.76	1.07 [-4.63 ; 7.11]	0.72
IL17a	-0.23 [-3.95 ; 3.63]	0.91	-4.20 [-9.76 ; 1.72]	0.16	3.19 [-1.17 ; 7.75]	0.15
<i>residential exposure</i>						
	Percentage change for 2 µg/m ³ [95%CI]	p trend	Percentage change for 5 µg/m ³ [95%CI]	p trend		
IL6	-0.39 [-4.68 ; 4.09]	0.86	-0.71 [-4.42 ; 3.16]	0.72		
IL8	-2.15 [-5.40 ; 1.21]	0.21	-2.03 [-4.86 ; 0.88]	0.17		
IL10	0.23 [-4.36 ; 5.05]	0.92	-0.60 [-4.58 ; 3.54]	0.77		
IL22	-2.13 [-9.82 ; 6.22]	0.61	-1.06 [-7.84 ; 6.22]	0.77		
TNF	-0.44 [-3.24 ; 2.45]	0.76	-0.53 [-2.96 ; 1.97]	0.68		
IL1	-3.77 [-10.83 ; 3.85]	0.32	-3.85 [-10.18 ; 2.93]	0.26		
IL17a	-0.067 [-5.48 ; 5.66]	0.98	-1.59 [-6.25 ; 3.30]	0.52		

Abbreviations: PM₁₀ = particulate matter with an aerodynamic diameter below 10 µm; NO₂ = nitrogen dioxide; O₃ = ozone

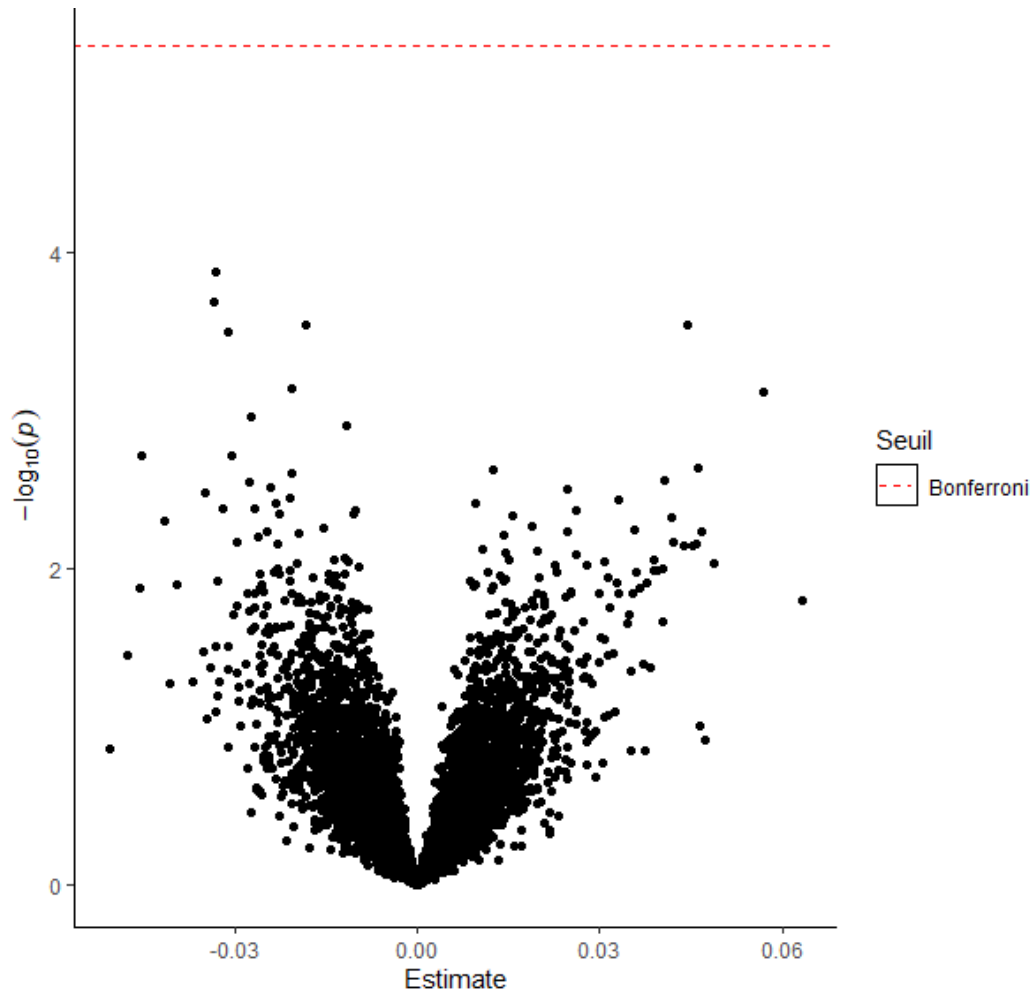
Annexe 3 : Résultats des analyses sur l'ensemble des métabolites (annotés et non annotés)



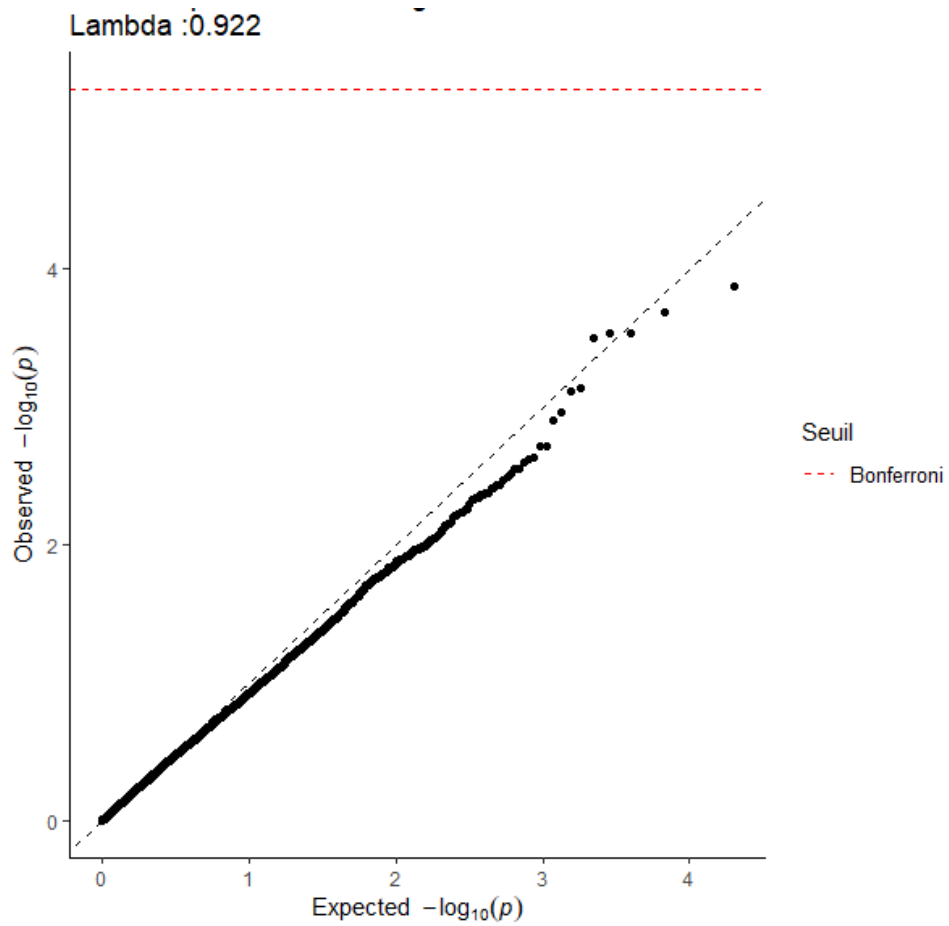
Volcano plot représentant les associations entre l'exposition à court terme aux PM_{10} et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation



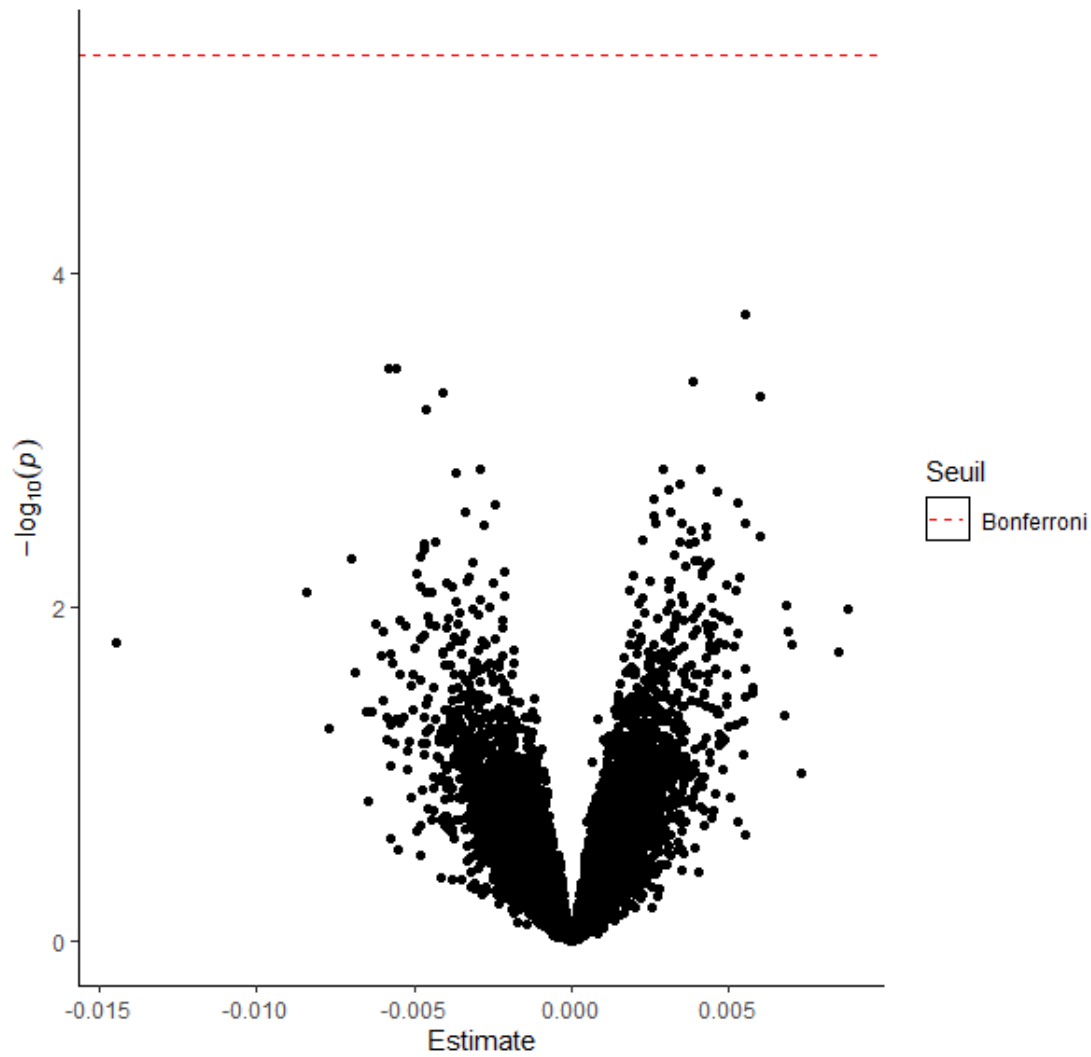
QQ plot représentant les associations entre l'exposition à court terme aux PM_{10} et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation



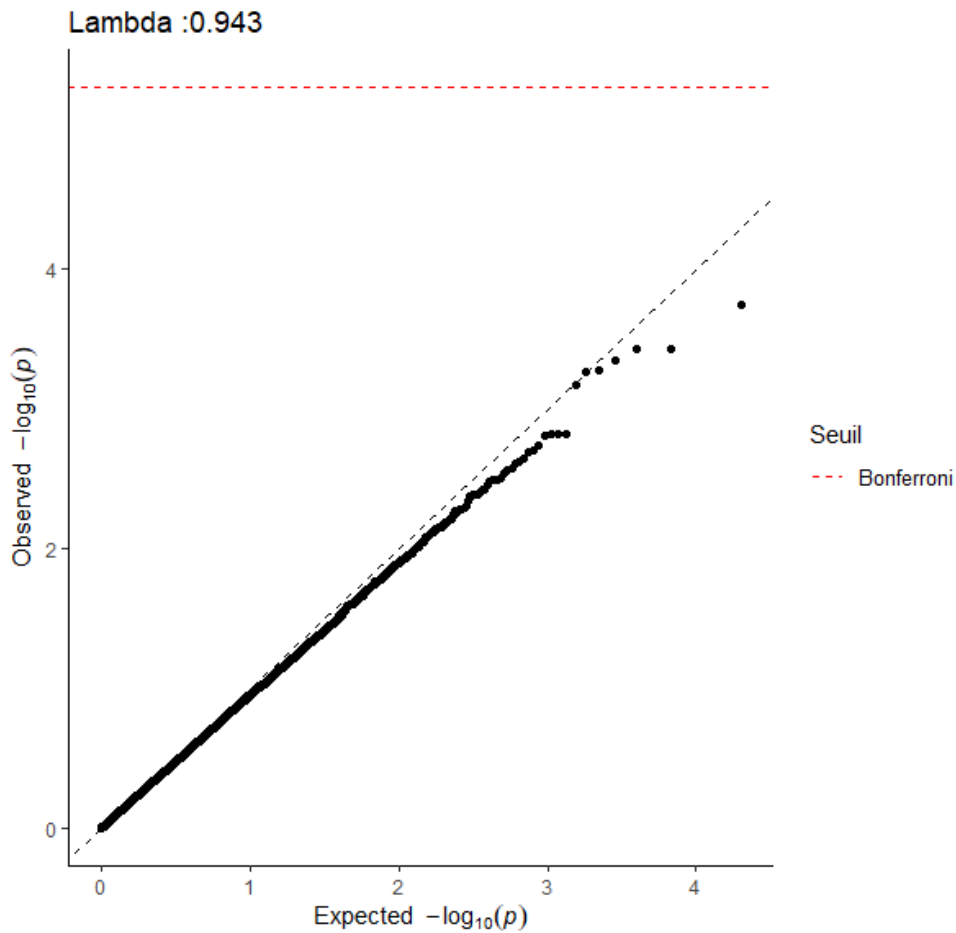
Volcano plot représentant les associations entre l'exposition résidentielle aux PM₁₀ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation



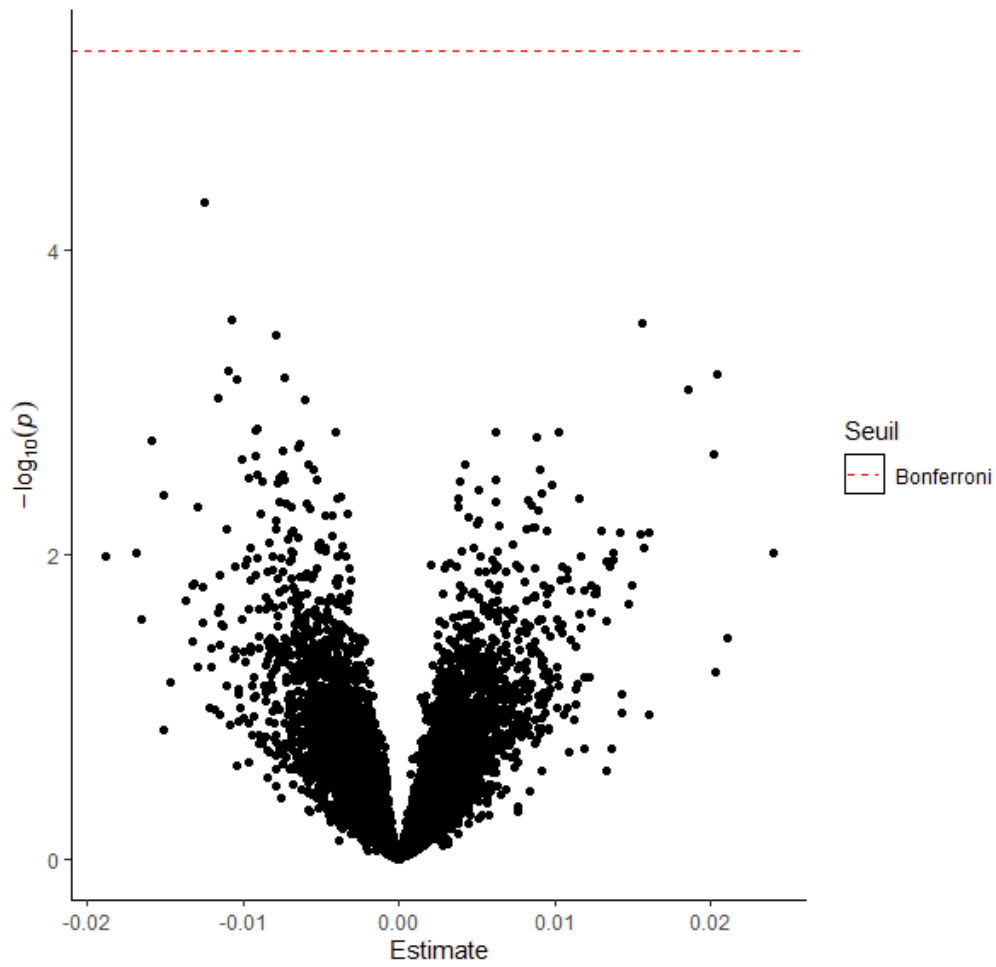
QQ plot représentant les associations entre l'exposition résidentielle aux PM_{10} et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation



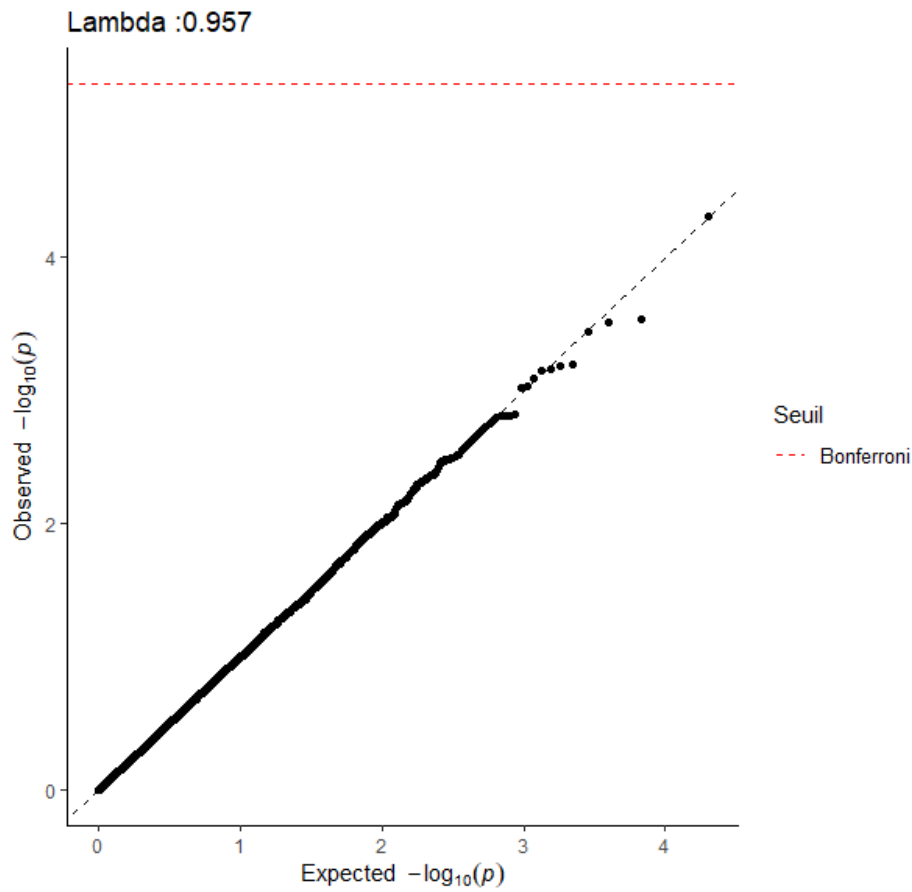
Volcano plot représentant les associations entre l'exposition à court terme au NO₂ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation



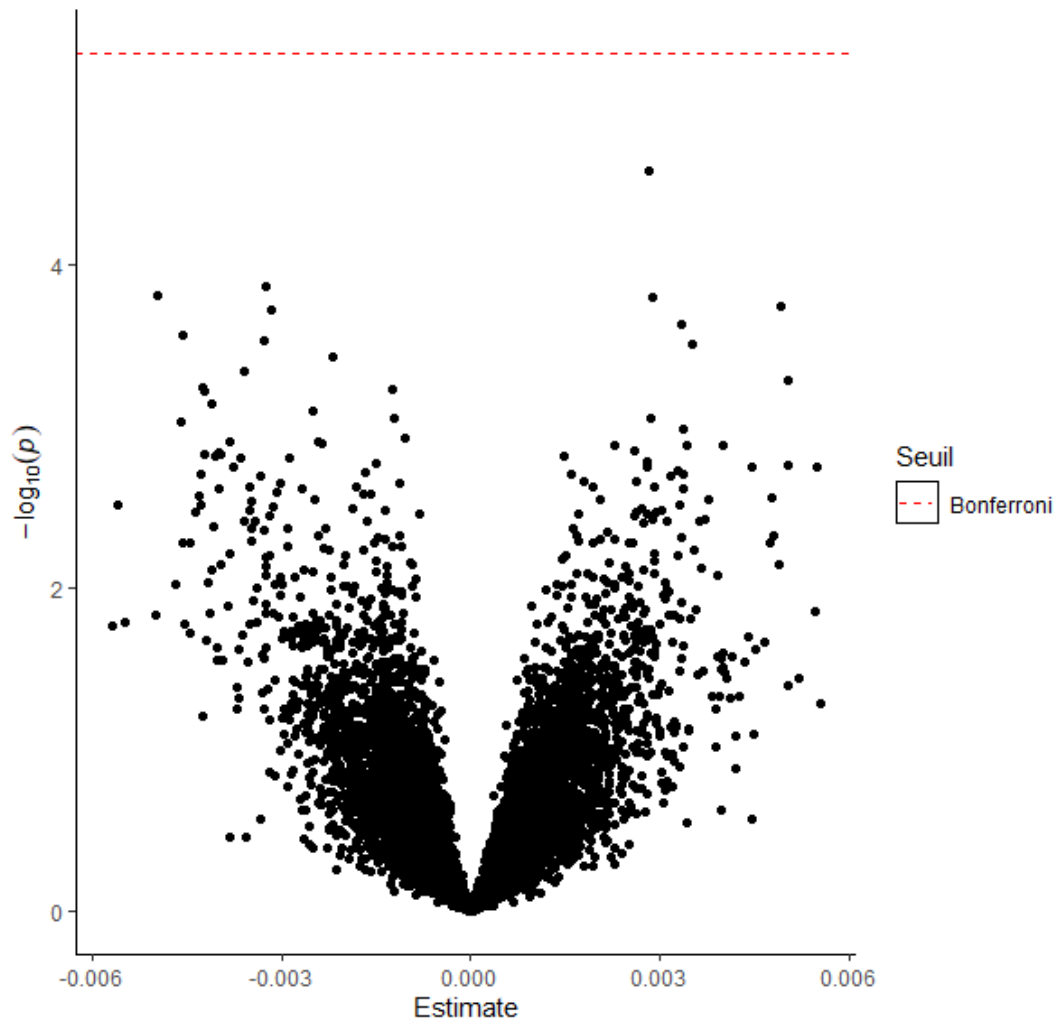
QQ plot représentant les associations entre l'exposition à court terme au NO_2 et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation



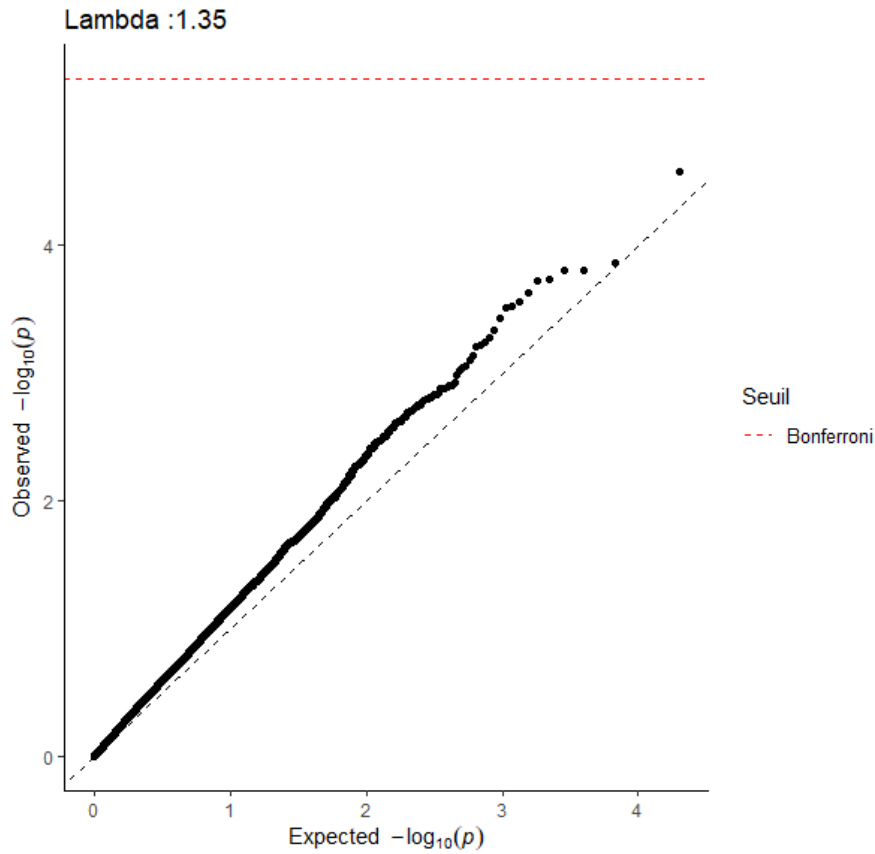
Volcano plot représentant les associations entre l'exposition résidentielle au NO₂ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation



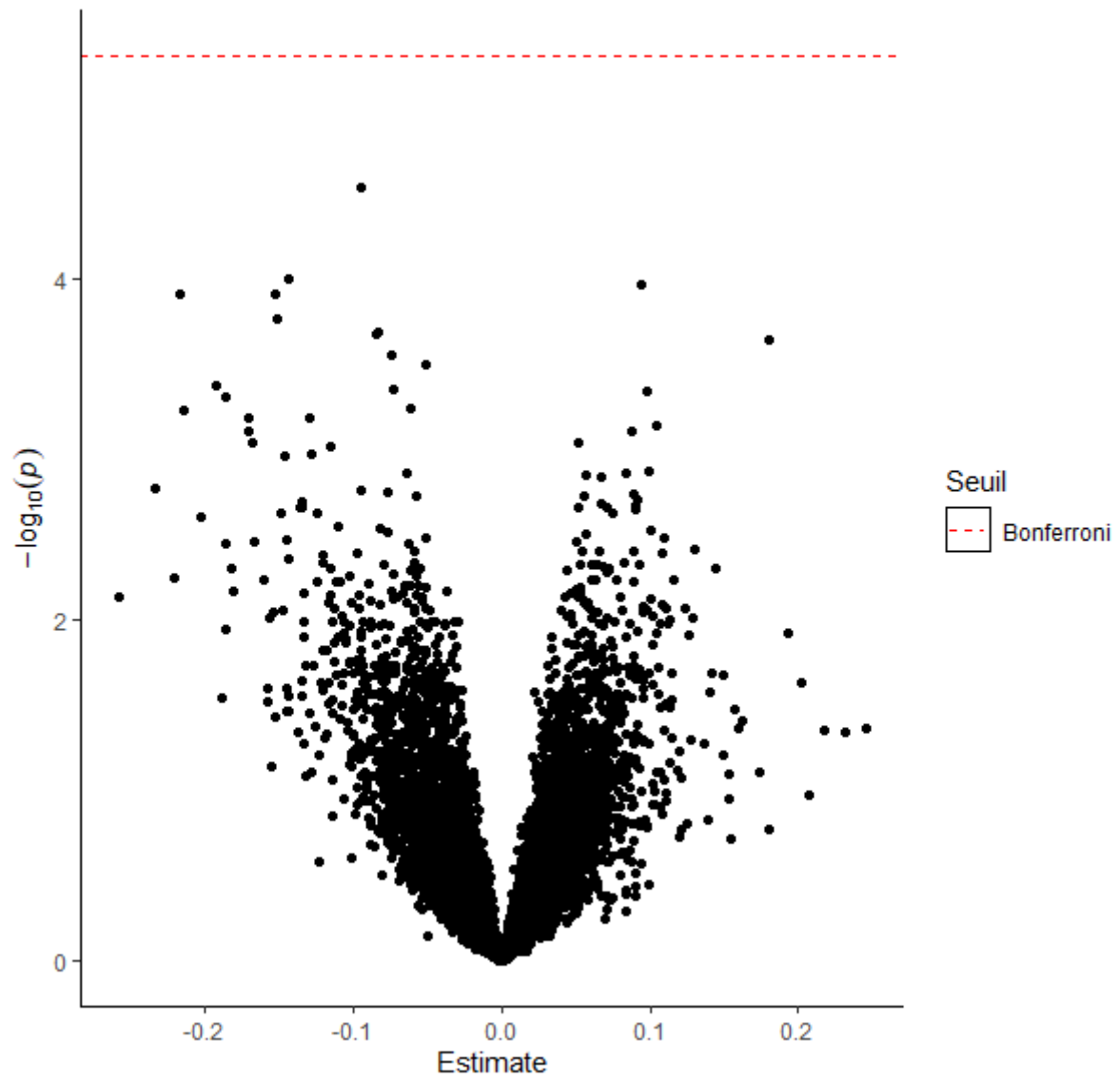
QQ plot représentant les associations entre l'exposition résidentielle au NO_2 et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC et le niveau d'éducation



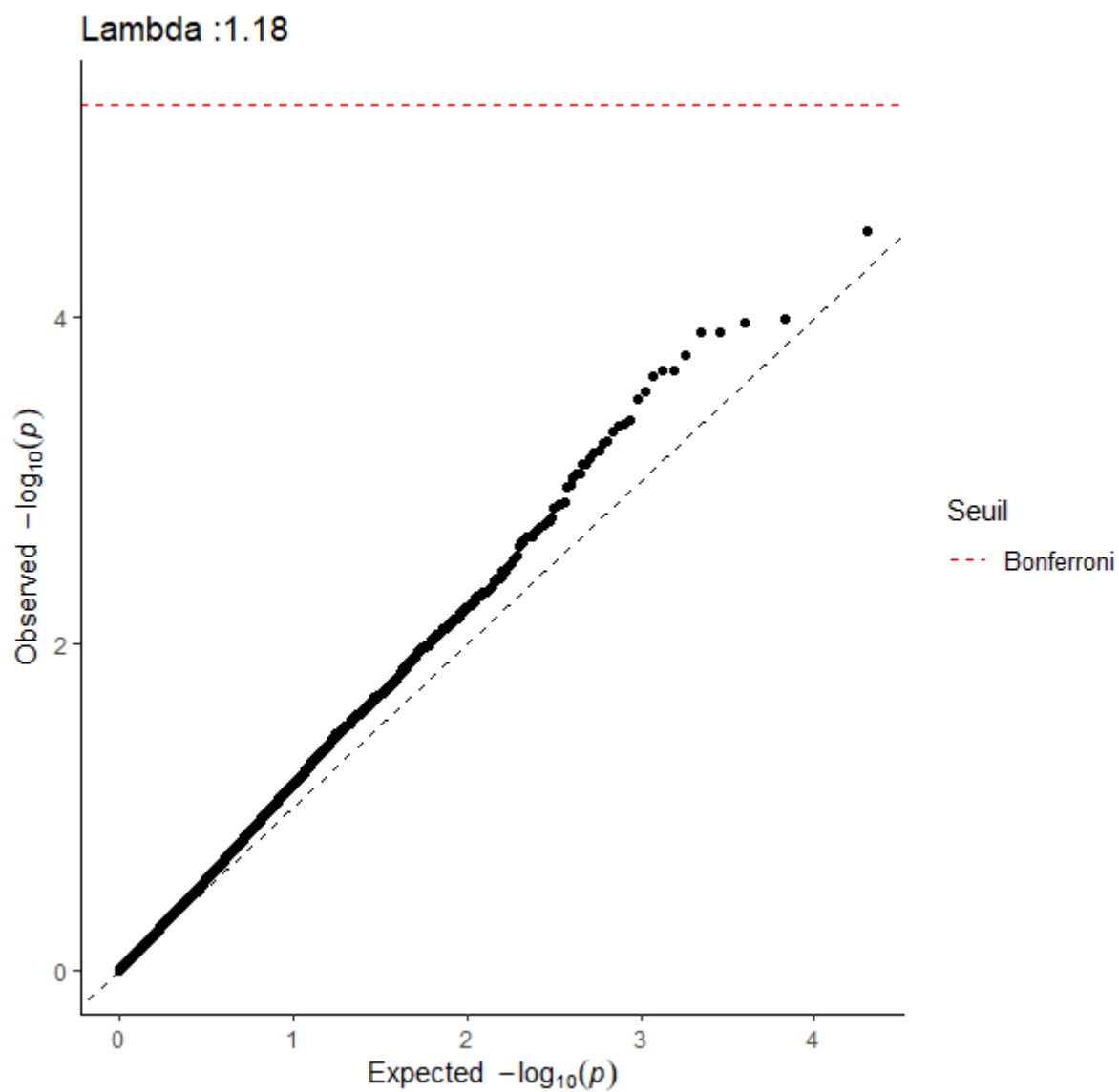
Volcano plot représentant les associations entre l'exposition à court terme à l'O₃ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation



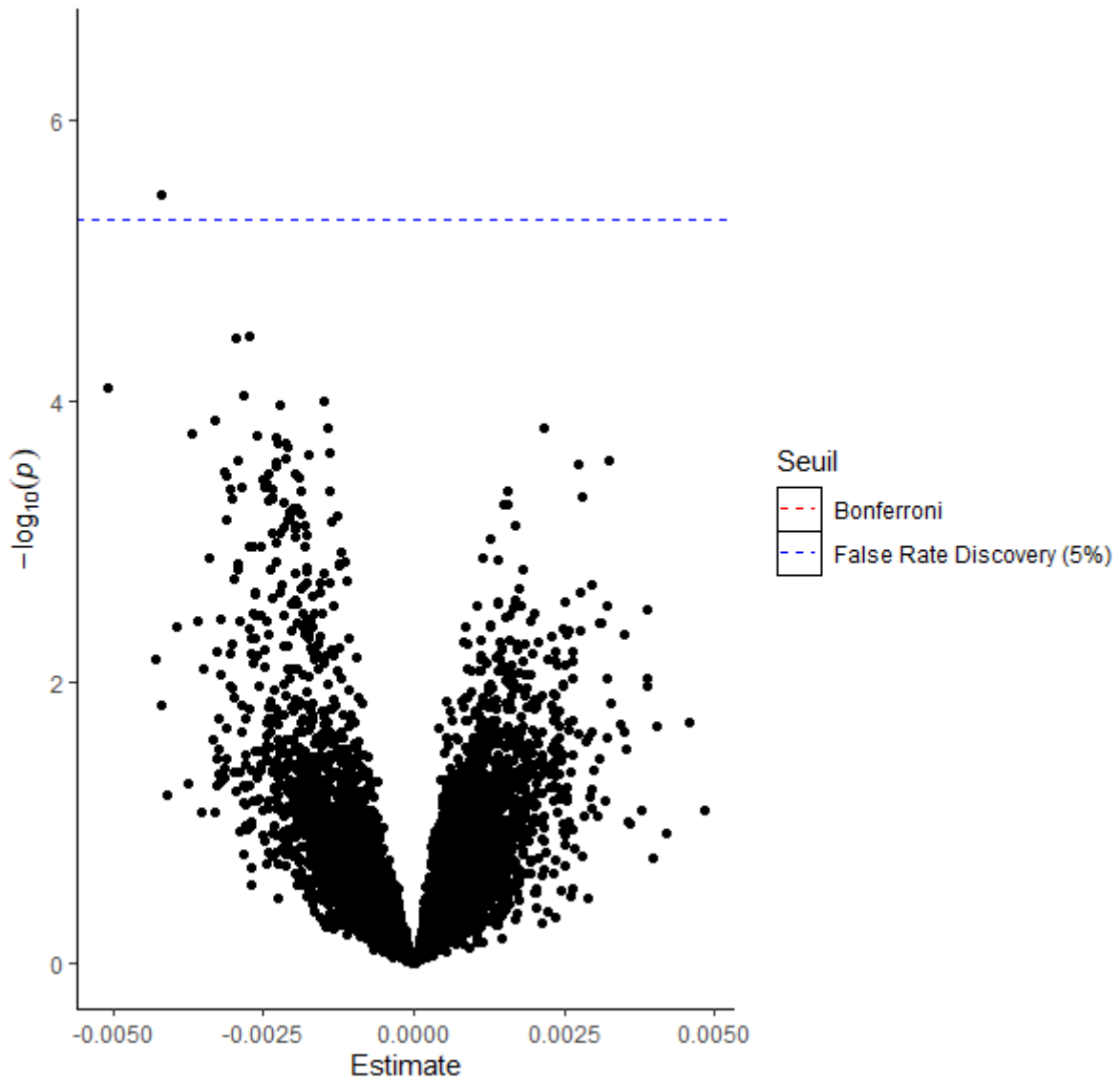
QQ plot représentant les associations entre l'exposition à court terme à l'O₃ et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine, l'année d'inclusion, l'IMC, les périodes de vacances scolaires, le jour de la semaine, le nombre de jours depuis le début de l'étude, la saison, la densité de population, l'humidité, les précipitations, la température, la pression atmosphérique et le niveau d'éducation



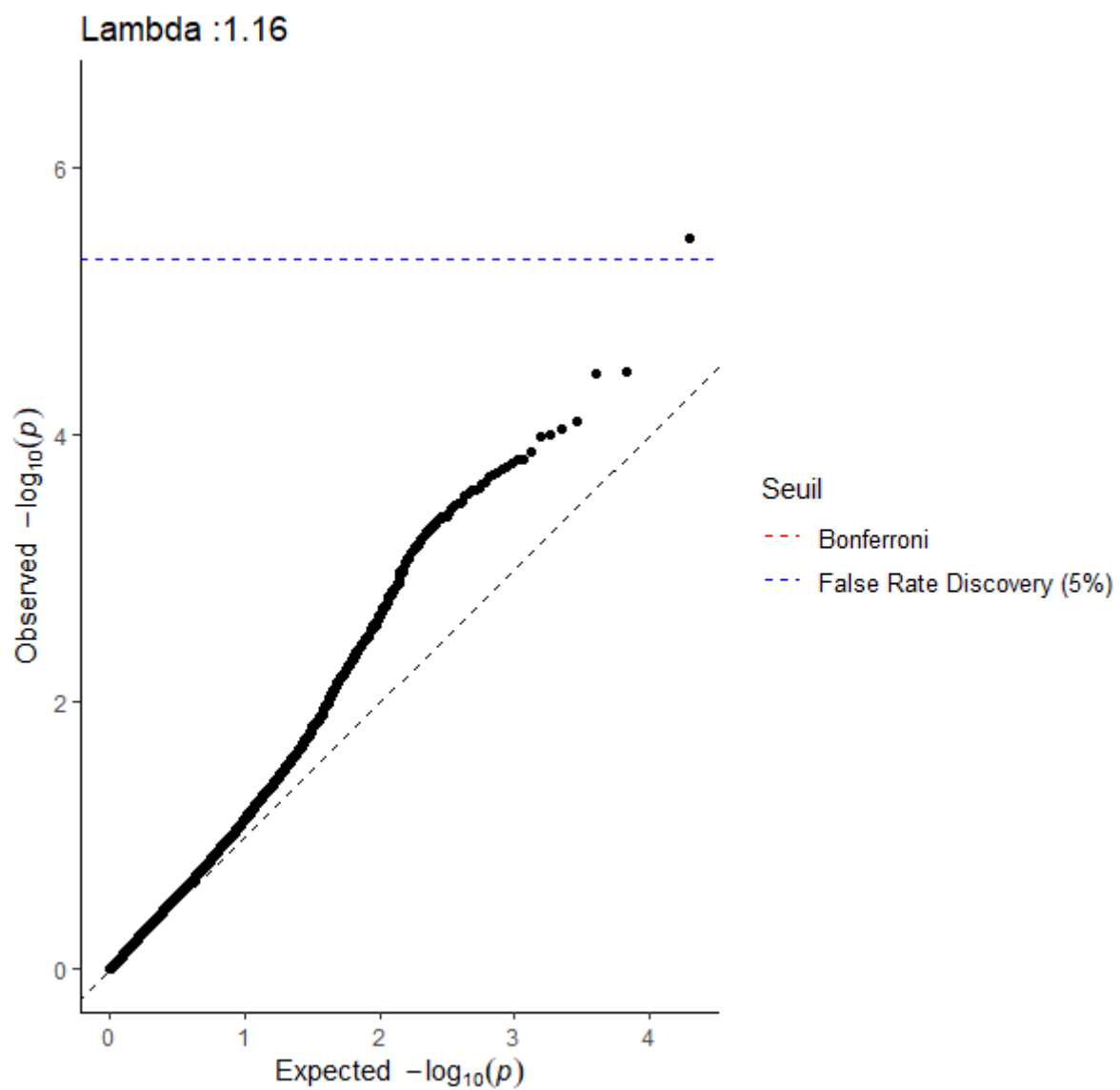
Volcano plot représentant les associations entre l'exposition aux moisissures et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



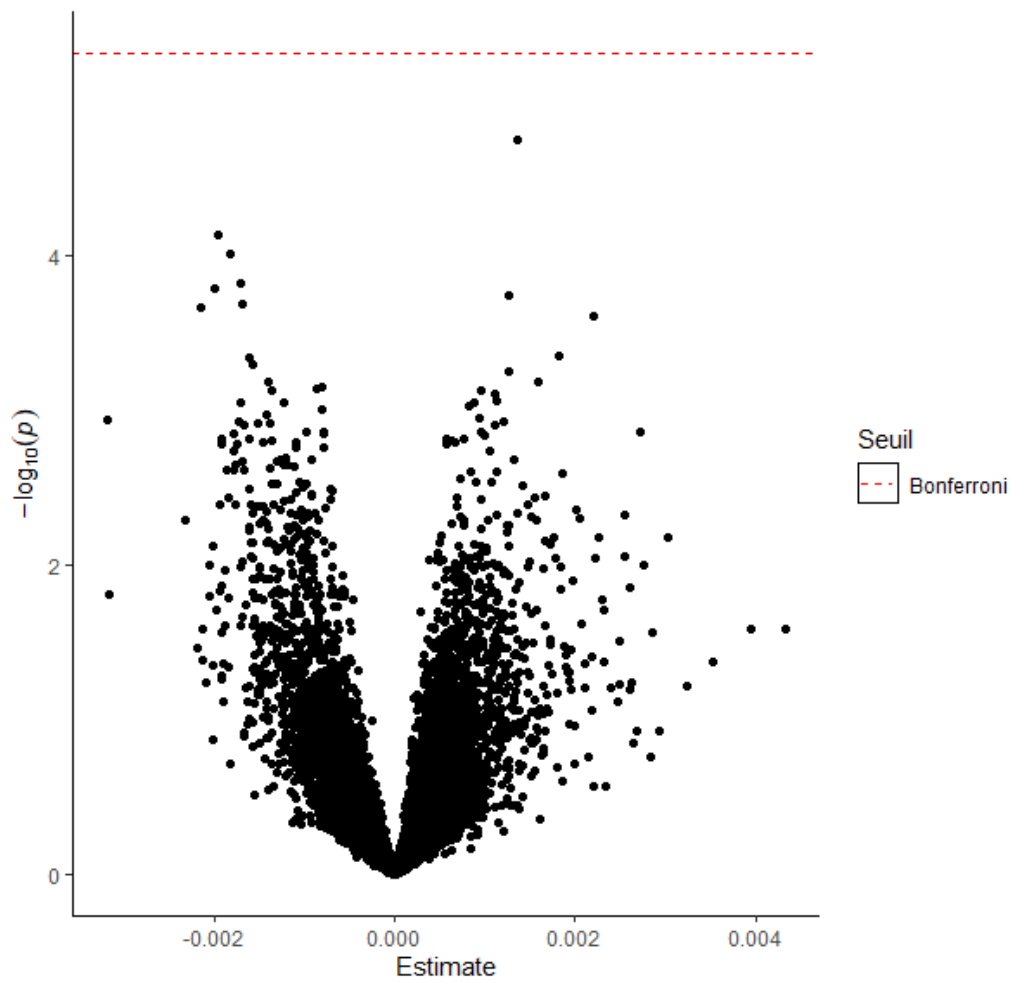
QQ plot représentant les associations entre l'exposition aux moisissures et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



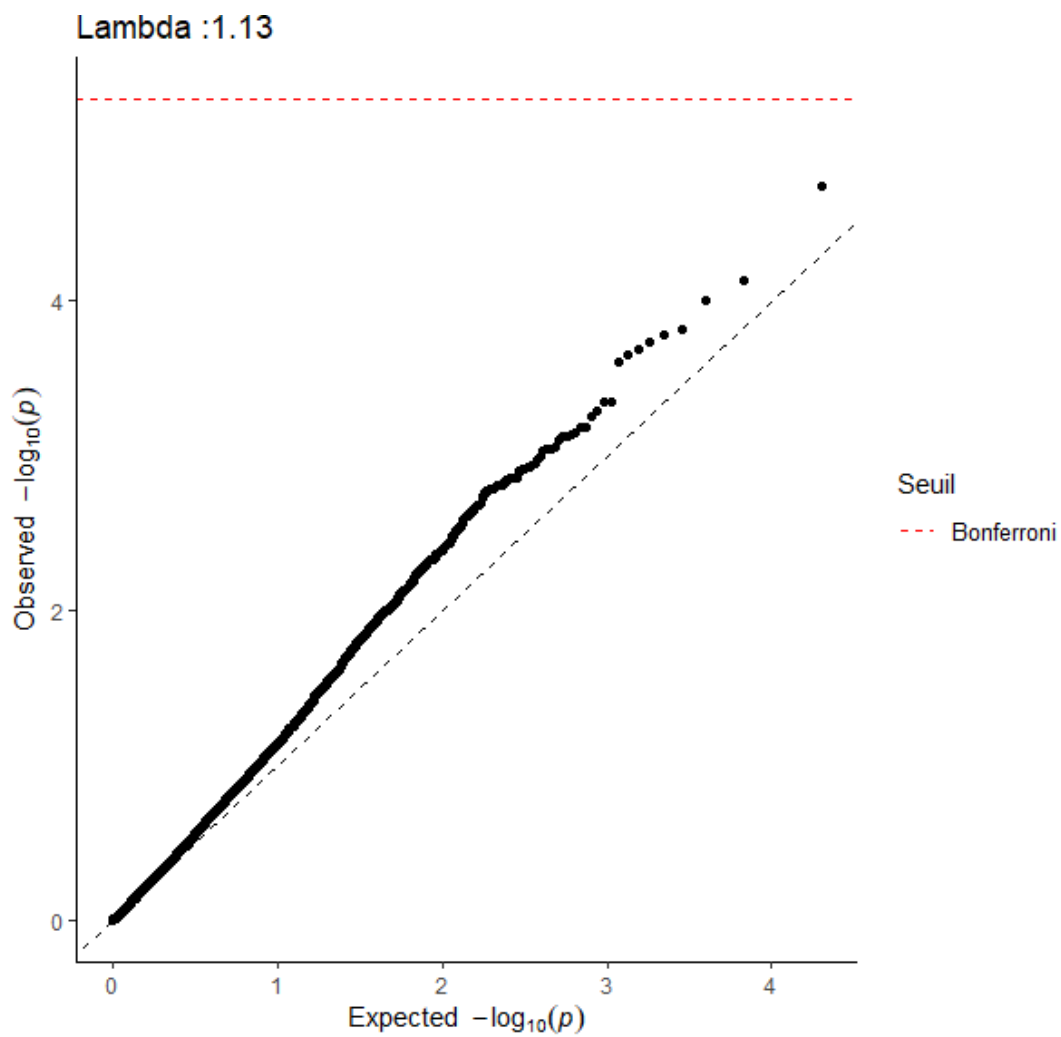
Volcano plot représentant les associations entre l'exposition aux vapeurs et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



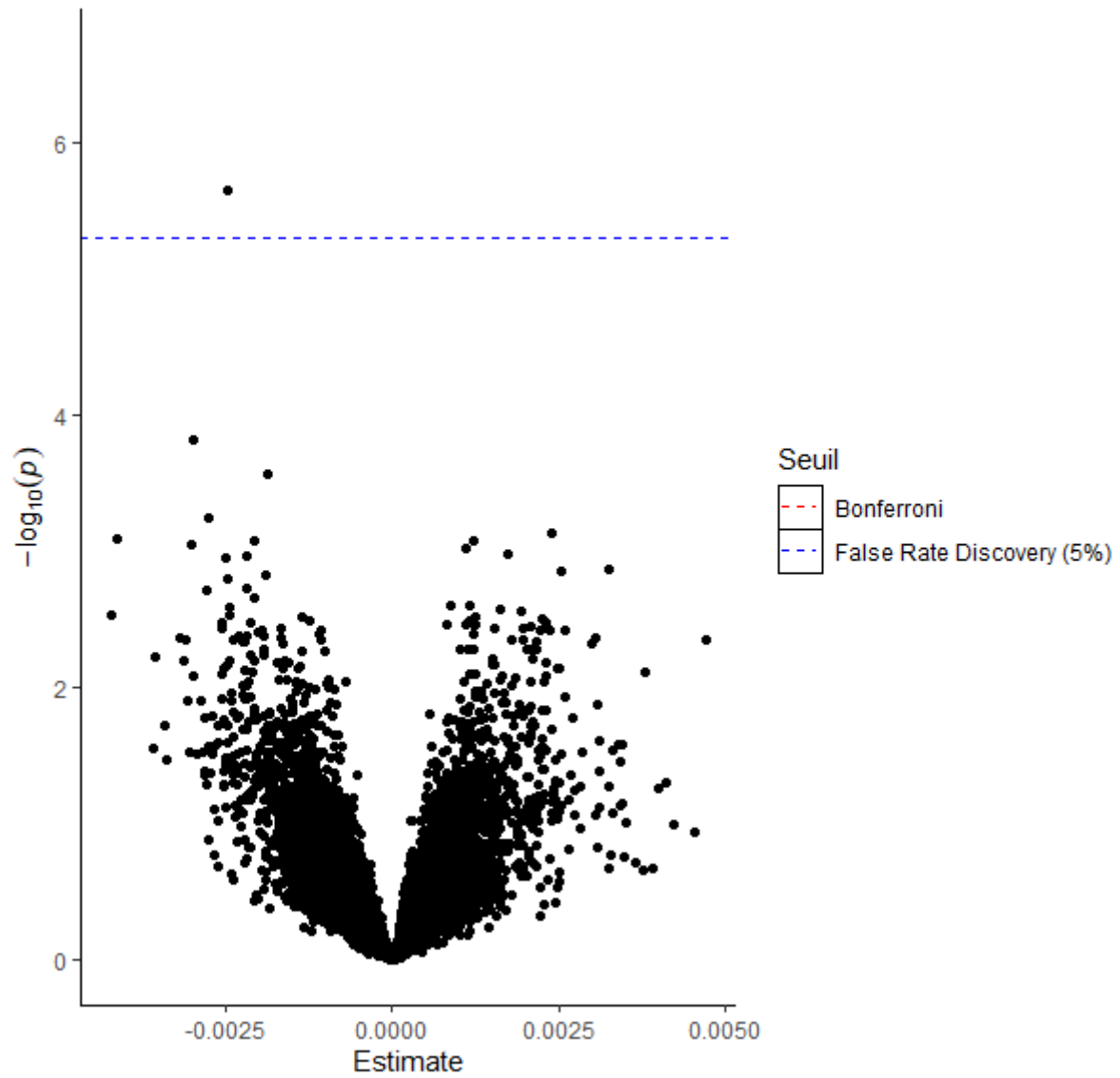
QQ plot représentant les associations entre l'exposition aux vapeurs et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



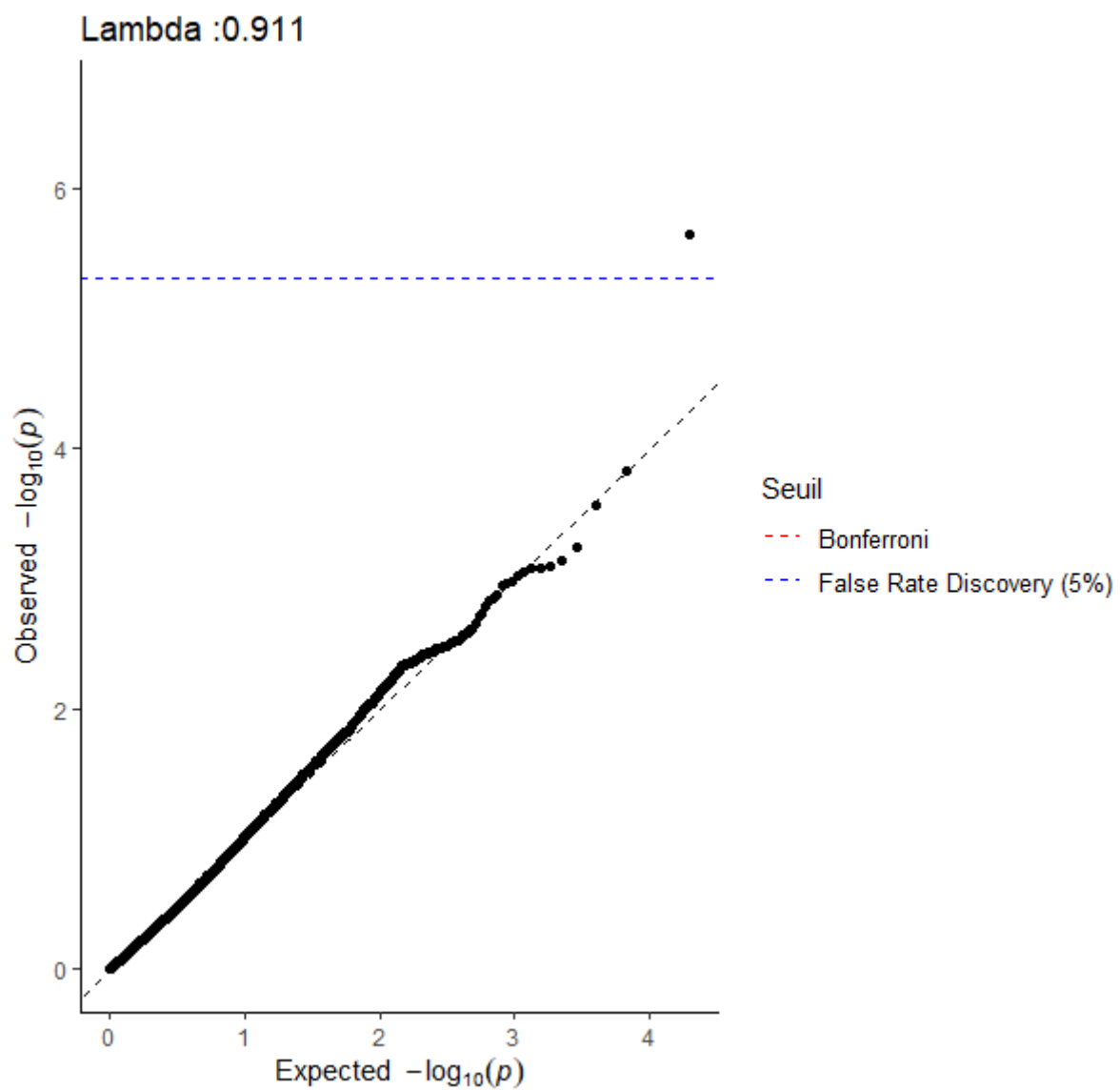
Volcano plot représentant les associations entre l'exposition aux poussières et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



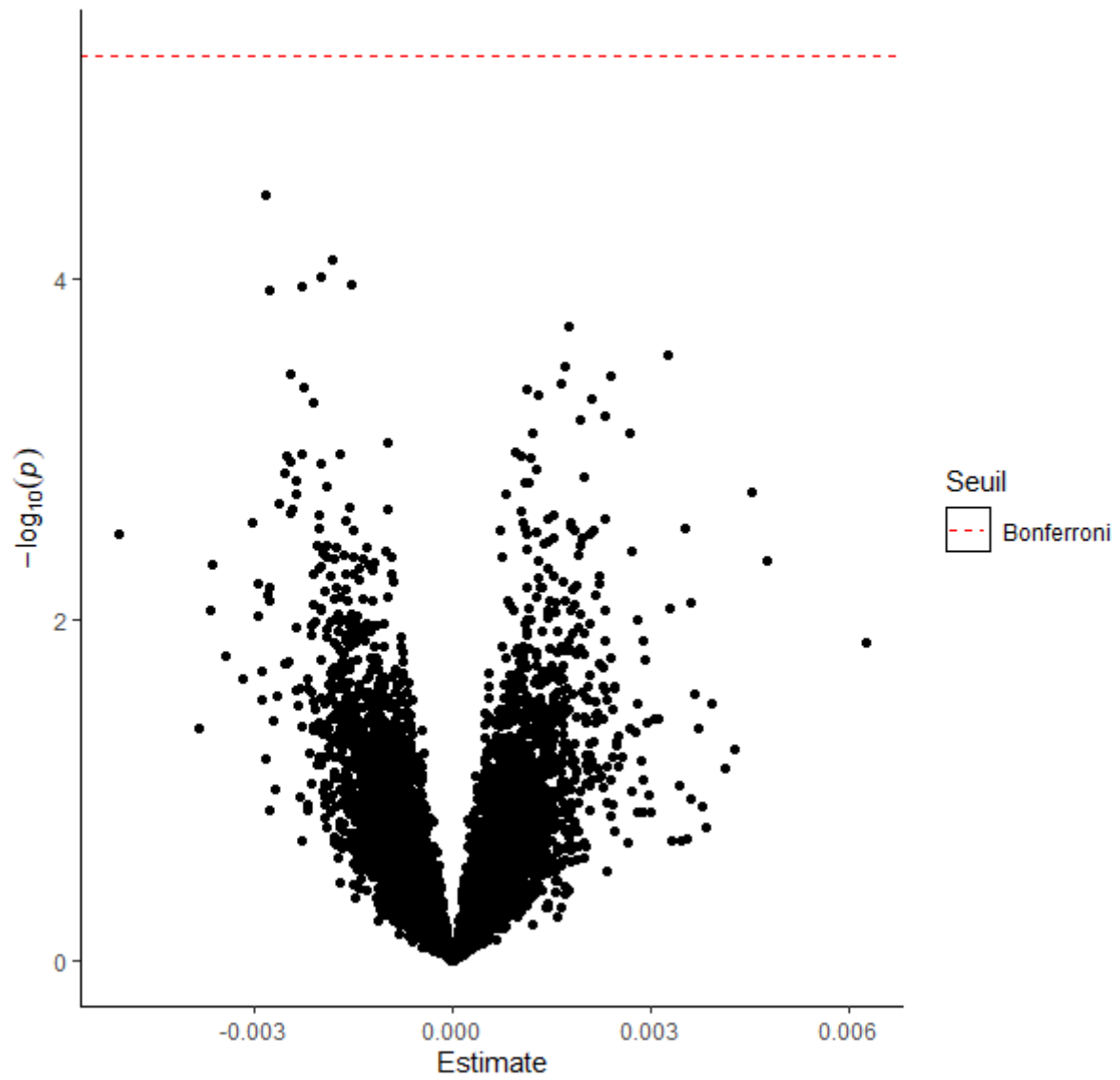
QQ plot représentant les associations entre l'exposition aux poussières et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



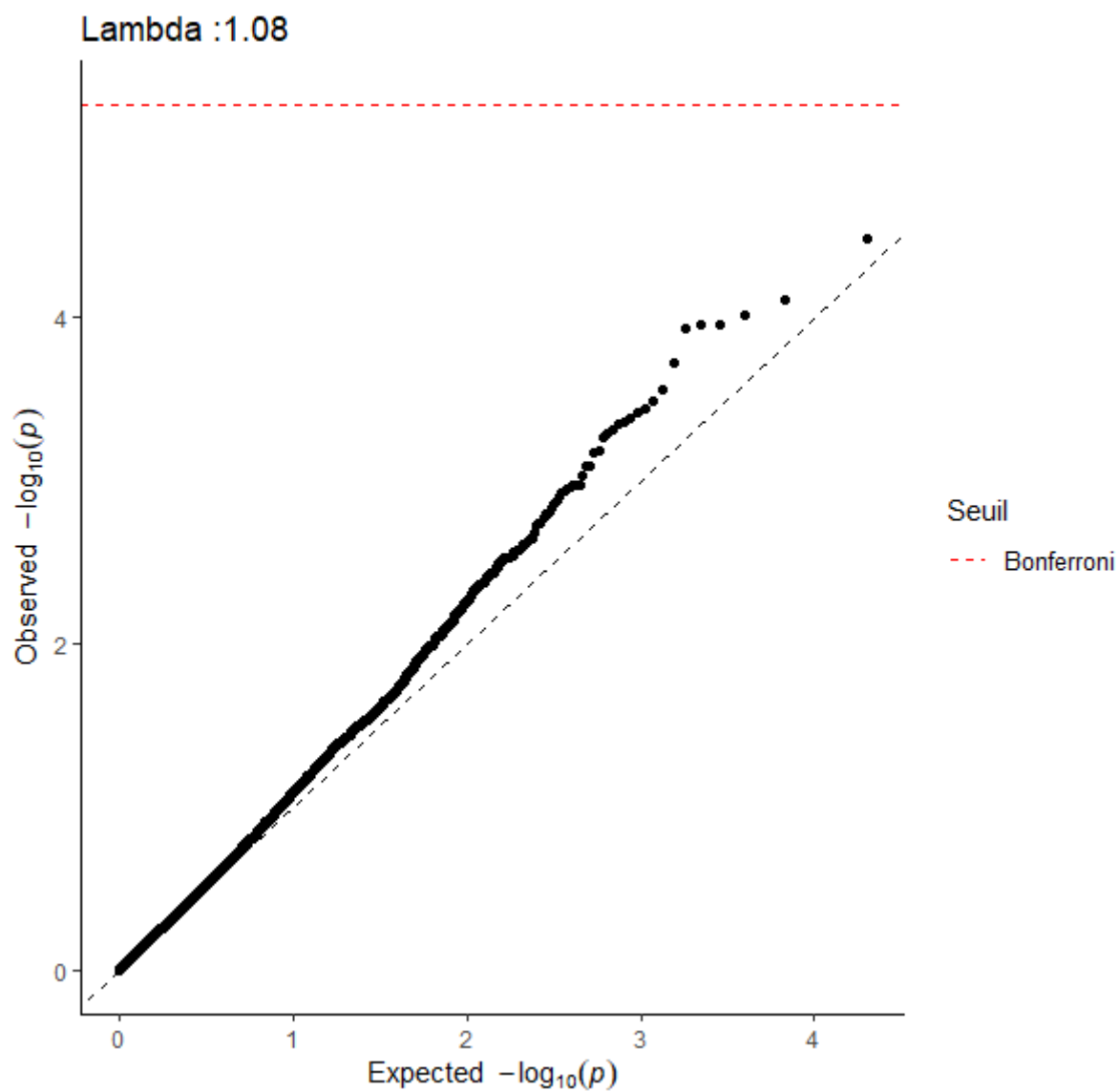
Volcano plot représentant les associations entre l'exposition aux gaz et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



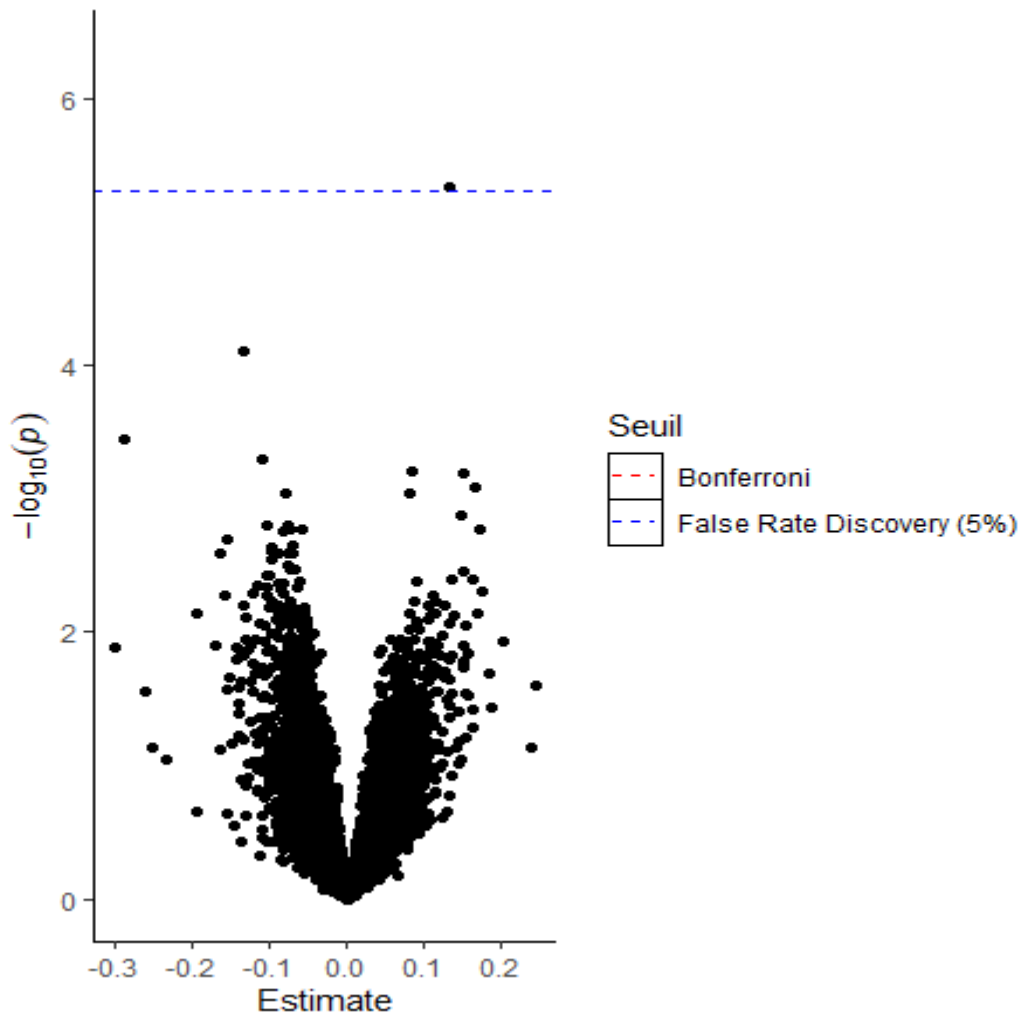
QQ plot représentant les associations entre l'exposition aux gaz et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



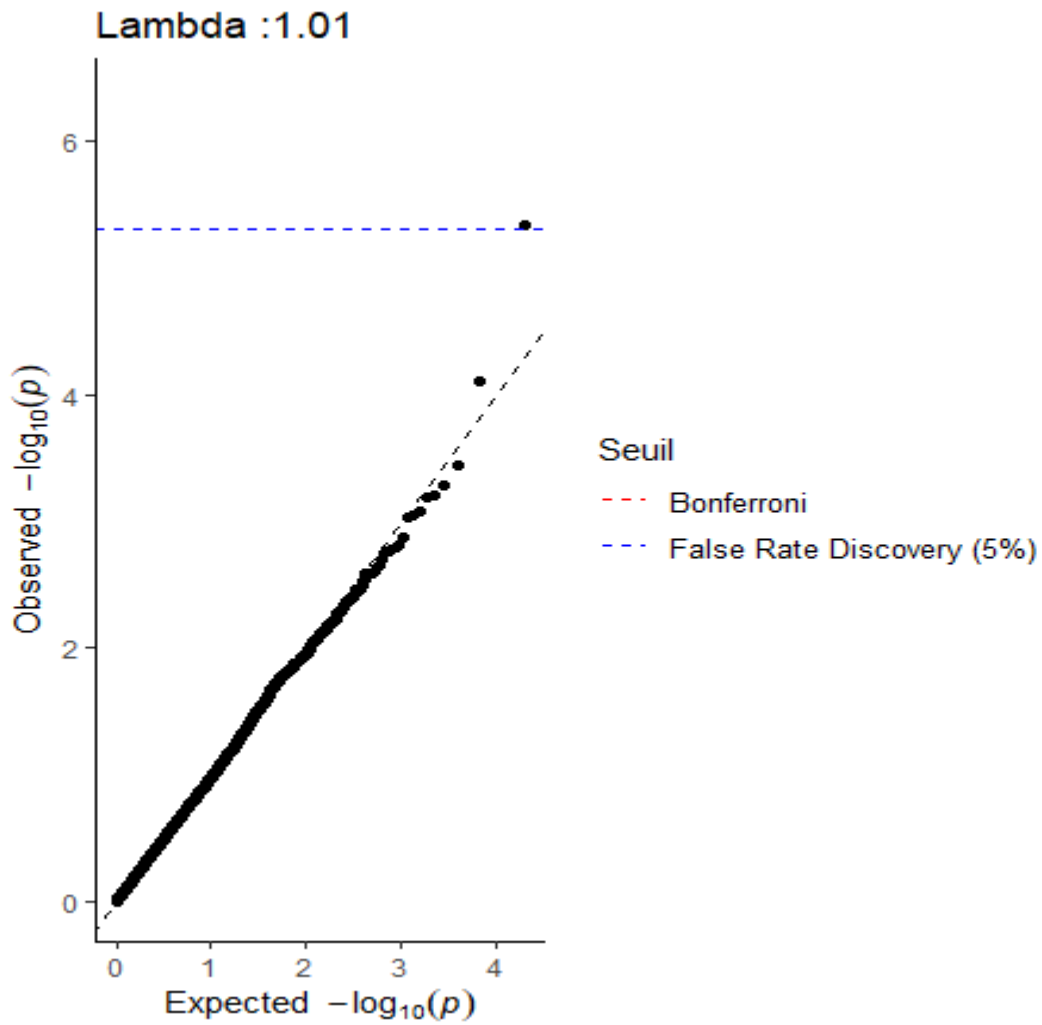
Volcano plot représentant les associations entre l'exposition aux fumées et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



QQ plot représentant les associations entre l'exposition aux fumées et les **métabolites**, ajustées sur l'âge, le sexe et la communauté urbaine



Volcano plot représentant les associations entre l'exposition au cadmium et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et lot d'analyse



QQ plot représentant les associations entre l'exposition au cadmium et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et lot d'analyse

Annexe 4 : Tableau indiquant le nombre de métabolites par voie métabolique

Tableau : nombre de métabolites par voie métabolique

Voie métabolique	Métabolites appartenant à la voie métabolomique
Métabolisme des folates	16015:Glutamic acid; 15428:Glycine
Métabolisme du fructose et du mannose	4208:Mannose; 2181:Fucose
Oxydation des acides gras	11060:Carnitine; 9300:Suberic acid; 15741:Succinic acid; 28867:Propionylcarnitine; 30915:Ketoglutaric acid; 18102:Octanoylcarnitine
Métabolisme des purines	17489:Cyclic AMP; 16335:Adenosine; 16704:Uridine; 17775:Uric acid; 15676:Allantoin; 17208:Ascorbic acid; 30837:Allantoic acid; 17596:Inosine; 17368:Hypoxanthine; 16708:Adenine; 18406:AICAR; 17712:Xanthine; 18107:Xanthosine; 15428:Glycine; 4208:Mannose; 16015:Glutamic acid; 32816:Pyruvic acid; 15873:Glucosamine phosphate
Métabolisme du pyruvate	32816:Pyruvic acid
Métabolisme des pyrimidines	17568:Uracil; 15901:Dihydrouracil; 15842:Orotidine; 16704:Uridine; 17562:Cytidine; 16040:Cytosine; 27468:Dihydrothymine; 17821:Thymine; 32816:Pyruvic acid; 16015:Glutamic acid; 18261:Ureidopropionic acid; 28816:Deoxyribose; 16742:Orotic acid
Biosynthèse des prostaglandines	15544:Prostaglandin E1; 17208:Ascorbic acid; 17242:Dehydroascorbic acid
Dégradation des protéines	16467:Arginine; 17196:Asparagine; 16015:Glutamic acid; 17295:Phenylalanine; 18019:Lysine; 16857:Threonine; 15971:Histidine; 17203:Proline; 17115:Serine; 15428:Glycine; 17895:Tyrosine; 16414:Valine; 17561:Cysteine
Métabolisme des lipides éther	17553:O-Phosphorylethanolamine
Métabolisme du propanoate	30831:Oxobutyric acid
Métabolisme des porphyrines	17208:Ascorbic acid; 17549:Aminolevulinic acid; 17381:Porphobilinogen; 17242:Dehydroascorbic acid
Métabolisme des peptides	16467:Arginine; 185298:Glycyl-leucine; 15428:Glycine; 17196:Asparagine; 17561:Cysteine; 17295:Phenylalanine; 16857:Threonine; 18019:Lysine; 15971:Histidine; 17115:Serine; 16015:Glutamic acid; 17203:Proline; 17895:Tyrosine; 16414:Valine
Voie des pentoses phosphates	28816:Deoxyribose; 15428:Glycine; 17295:Phenylalanine; 33198:Gluconic acid
Biosynthèse de la phénylalanine, de la tyrosine et du tryptophane	30915:Ketoglutaric acid; 17895:Tyrosine; 16015:Glutamic acid; 18344:Kynurenic acid; 16946:Kynurenine; 15793:Hydroxyanthranilic acid; 18243:Dopamine; 17208:Ascorbic acid; 16765:Tryptamine; 16411:Indoleacetic acid; 30851:Phenylpyruvic acid; 17295:Phenylalanine; 15428:Glycine; 217069:Xanthurenic acid; 17561:Cysteine; 1127735:Vanillomandelic acid; 32816:Pyruvic acid; 17242:Dehydroascorbic acid; 4208:Mannose; 27480:Phenylacetylglutamine; 17884:Phenylacetylglutamine
Métabolisme de la phénylalanine	15428:Glycine; 30915:Ketoglutaric acid; 17295:Phenylalanine; 30851:Phenylpyruvic acid; 27480:Phenylacetylglutamine; 16015:Glutamic acid
Métabolisme de la cystéine et de la méthionine	17115:Serine; 17561:Cysteine; 16015:Glutamic acid; 32816:Pyruvic acid; 16283:Cystine; 30915:Ketoglutaric acid; 27891:Cysteine S-sulfate; 15891:Taurine; 30831:Oxobutyric acid; 16335:Adenosine
Métabolisme des éicosanoïdes	15428:Glycine; 16015:Glutamic acid; 15544:Prostaglandin E1
Interconversions des pentoses et du glucuronate	18403:Arabitol
Métabolisme des médicaments	32636:Acetaminophen glucuronide; 46195:Acetaminophen (Acetamidophenol); 30915:Ketoglutaric acid; 17775:Uric acid; 15891:Taurine; 15428:Glycine; 16015:Glutamic acid
Biosynthèse de la pantothéine et de la coenzyme A	7916:Pantothenic acid; 17561:Cysteine
Synthèse de la coenzyme A	17561:Cysteine; 7916:Pantothenic acid
Métabolisme du tryptophane	30915:Ketoglutaric acid; 16946:Kynurenine; 16675:Quinolinic acid; 217069:Xanthurenic acid; 16015:Glutamic acid; 18344:Kynurenic acid; 16411:Indoleacetic acid; 27823:Hydroxyindoleacetic acid
Métabolisme du cholestérol	25351:Mevalonic acid

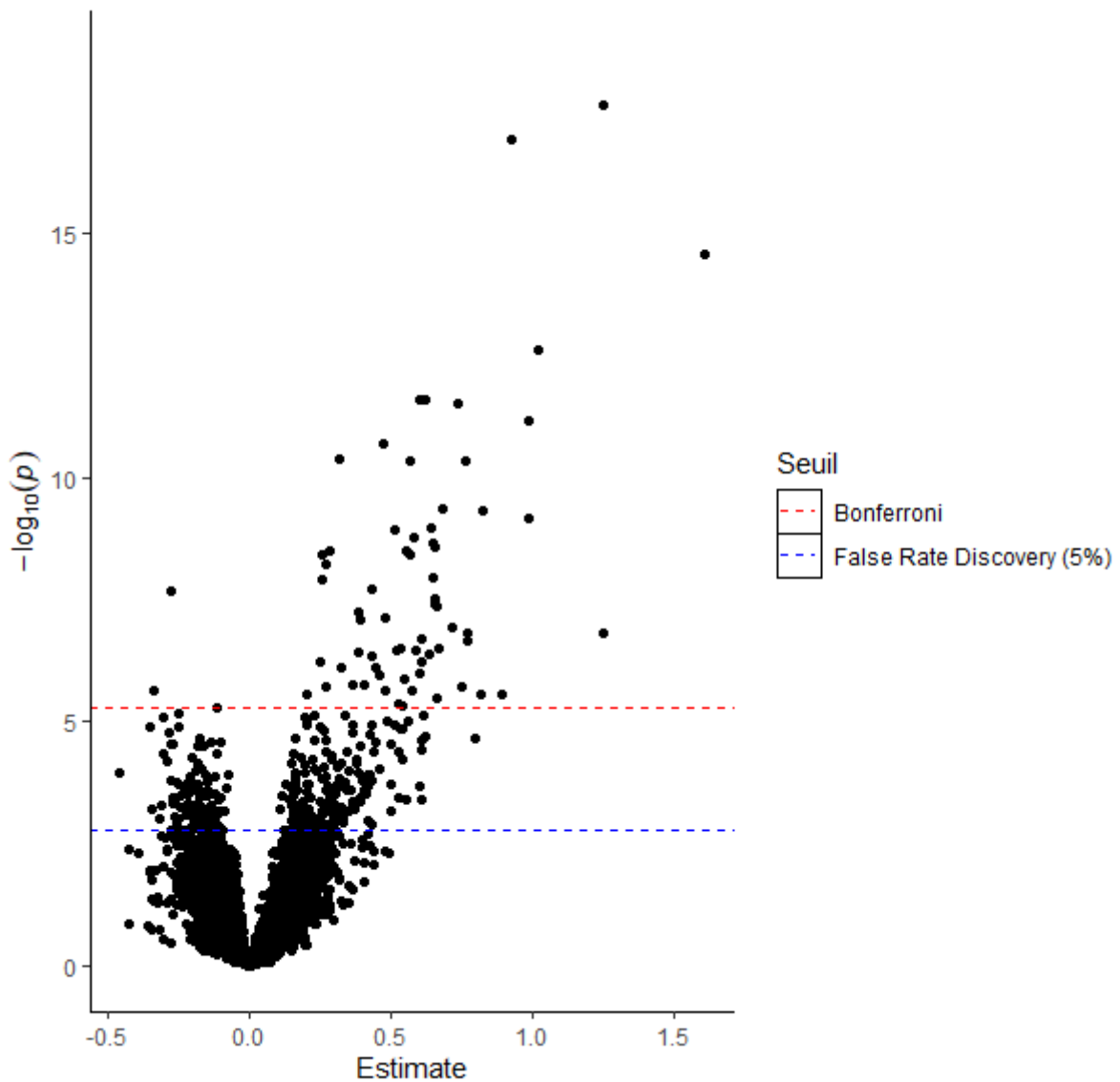
Tableau : nombre de métabolites par voie métabolique (suite)

Voie métabolique	Métabolites appartenant à la voie métabolomique
Métabolisme de la vitamine C	18268:Glucuronic acid gamma-lactone; 16002:Saccharic acid; 17242:Dehydroascorbic acid
Métabolisme de la vitamine B6	17310:Pyridoxal; 16709:Pyridoxine; 17405:Pyridoxic acid
Navette de la carnitine mitochondriale	11060:Carnitine; 18102:Octanoylcarnitine; 28867:Propionylcarnitine
Navette de la carnitine du réticulum endoplasmique	11060:Carnitine
Métabolisme de la tyrosine	18243:Dopamine; 17895:Tyrosine; 30915:Ketoglutaric acid; 17561:Cysteine; 16424:Phenylacetaldehyde; 17295:Phenylalanine; 16015:Glutamic acid; 15428:Glycine
Biosynthèse du cholestérol 1 : voie de Bloch	25351:Mevalonic acid
Métabolisme de la valine, de la leucine et de l'isoleucine	30915:Ketoglutaric acid; 15428:Glycine; 16015:Glutamic acid; 16414:Valine
Navette de la carnitine peroxysomale	11060:Carnitine; 28867:Propionylcarnitine; 18102:Octanoylcarnitine
Métabolisme du butanoate	30915:Ketoglutaric acid; 30831:Oxobutyric acid
Navette de la carnitine cytosolique	18102:Octanoylcarnitine; 11060:Carnitine; 28867:Propionylcarnitine
Métabolisme des acides dibasiques ramifiés en C5	16600:Mesaconic acid; 30838:Itaconic acid; 32816:Pyruvic acid
Cycle de l'acide tricarboxylique et métabolisme du glyoxylate et des dicarboxylates	30915:Ketoglutaric acid; 30769:Citric acid; 15741:Succinic acid; 32398:Glyceric acid; 16995:Oxalic acid
Métabolisme de la biopiptéine	64277:Dihydrobiopterin
Biosynthèse du squelette des terpénoïdes	17561:Cysteine
Métabolisme du soufre	30915:Ketoglutaric acid; 17724:N,N-Dimethylglycine; 16015:Glutamic acid; 32816:Pyruvic acid
Métabolisme de la biotine	18019:Lysine
Métabolisme de la thiamine	18385:Thiamine; 32816:Pyruvic acid
Métabolisme des sphingolipides	17115:Serine; 17553:O-Phosphorylethanolamine; 45744:N-Acetylneuraminic acid
Biosynthèse de la sérotonine et de la mélatonine	27823:Hydroxyindoleacetic acid; 17697:N-Acetyl-serotonin
Biosynthèse des acides biliaires	15891:Taurine; 15428:Glycine; 17561:Cysteine
Métabolisme des stéroïdes	17895:Tyrosine; 27584:Aldosterone
Métabolisme de la bêta-alanine	16015:Glutamic acid; 18261:Ureidopropionic acid; 16610:Spermidine; 15971:Histidine; 15727:Carnosine; 30915:Ketoglutaric acid; 16865:γ-aminobutyric acid
β-oxydation de l'acide phytanique (peroxysomale)	30915:Ketoglutaric acid; 15741:Succinic acid
Métabolisme des acides gras oméga-3	16015:Glutamic acid

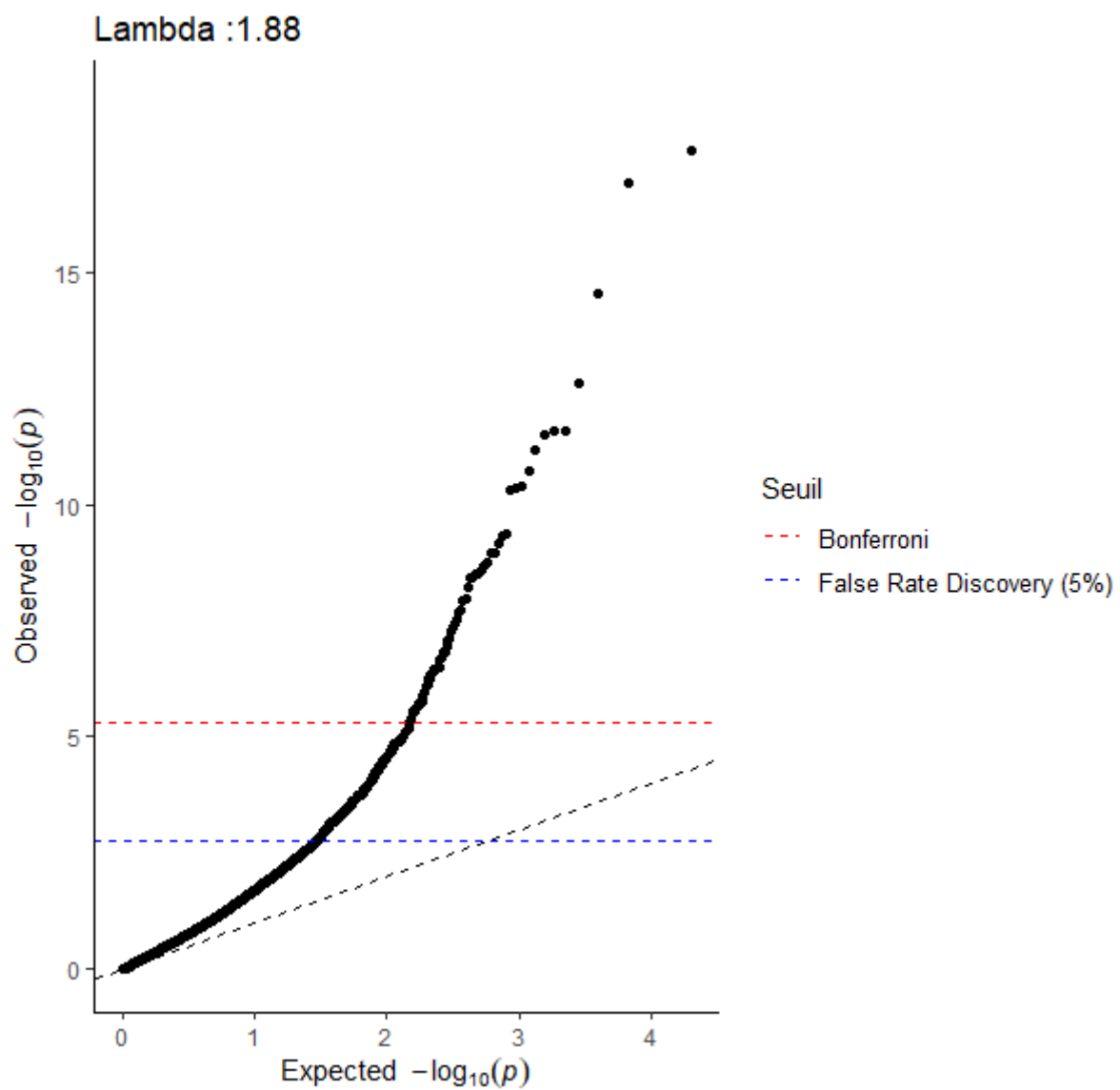
Tableau : nombre de métabolites par voie métabolique (fin)

Voie métabolique	Métabolites appartenant à la voie métabolomique
Métabolisme des O-glycanes Métabolisme de l'ascorbate et des aldarates	17115:Serine 17208:Ascorbic acid; 17242:Dehydroascorbic acid; 18268:Glucuronic acid gamma-lactone; 18019:Lysine; 16995:Oxalic acid; 16002:Saccharic acid 17208:Ascorbic acid
Métabolisme de la vitamine E Métabolisme des N-glycanes Diverses voies	45744:N-Acetylneuraminic acid; 2181:Fucose; 4208:Mannose 15978:Glycerol phosphate; 17295:Phenylalanine; 32816:Pyruvic acid; 1127735:Vanillomandelic acid
Métabolisme des nucléotides	15901:Dihydrouracil; 17562:Cytidine; 16335:Adenosine; 17596:Inosine; 17489:Cyclic AMP; 17568:Uracil; 18107:Xanthosine; 16704:Uridine; 16708:Adenine
Métabolisme de l'arginine et de la proline	15729:Ornithine; 16467:Arginine; 16610:Spermidine; 17509:Deoxymethylthioadenosine; 17203:Proline; 17768:N-Acetyl-putrescine; 16015:Glutamic acid; 30915:Ketoglutaric acid; 16349:Citrulline; 18095:Hydroxyprolines; 16919:Creatine; 18019:Lysine; 16708:Adenine; 17645:Acetamidobutanoic acid; 12575:N-Acetyl-glutamic acids; 15428:Glycine; 16737:Creatinine; 15741:Succinic acid; 16865:γ-aminobutyric acid; 32816:Pyruvic acid; 16344:Guanidineacetic acid
Métabolisme du nicotinate et de la nicotinamide	16675:Quinolinic acid; 15793:Hydroxyanthranilic acid; 16015:Glutamic acid; 30915:Ketoglutaric acid
Métabolisme de l'alanine, de l'aspartate et du glutamate	16265:Succinic semialdehyde; 32816:Pyruvic acid; 30915:Ketoglutaric acid; 17196:Asparagine; 17895:Tyrosine; 21547:N-Acetyl-aspartic acid; 16015:Glutamic acid; 16467:Arginine; 16349:Citrulline; 16865:γ-aminobutyric acid; 15741:Succinic acid; 30831:Oxobutyric acid; 15682:Argininosuccinic acid
Biosynthèse des aminoacyl-ARNt	16857:Threonine; 17203:Proline; 17115:Serine; 17295:Phenylalanine; 18019:Lysine; 16414:Valine; 16015:Glutamic acid; 15428:Glycine; 17561:Cysteine; 15971:Histidine; 17895:Tyrosine; 17196:Asparagine; 16467:Arginine
Métabolisme des autres acides aminés	32816:Pyruvic acid; 30831:Oxobutyric acid; 16335:Adenosine; 17115:Serine; 15891:Taurine
Métabolisme des sucres aminés et des nucléotides-oses	15873:Glucosamine phosphate; 32816:Pyruvic acid; 45744:N-Acetylneuraminic acid; 16015:Glutamic acid
Métabolisme de la lysine	30915:Ketoglutaric acid; 18019:Lysine; 17311:N6,N6,N6-Trimethyl-lysine; 16927:Saccharopine; 16015:Glutamic acid; 15741:Succinic acid; 11060:Carnitine; 15428:Glycine; 37024:Amino adipic acid
Métabolisme des acylglycérides	15978:Glycerol phosphate
Dégradation du kératine sulfate	45744:N-Acetylneuraminic acid; 2181:Fucose
Métabolisme du linoléate	48131:Azelaic acid; 17208:Ascorbic acid
Métabolisme des leucotriènes	15428:Glycine; 16015:Glutamic acid
Métabolisme de l'histidine	30817:Imidazoleacrylic acids (urocanic acid); 15971:Histidine; 16015:Glutamic acid; 16865:γ-aminobutyric acid; 15428:Glycine; 16974:Imidazoleacetic acid; 15727:Carnosine
Métabolisme de l'hippurate	15428:Glycine; 30746:Benzoic acid; 18089:Hippuric acid
Voies isolées	18040:Hydroxylysine; 18089:Hippuric acid; 15573:Aminoethylphosphonic acid; 16015:Glutamic acid; 17242:Dehydroascorbic acid; 30915:Ketoglutaric acid; 17208:Ascorbic acid; 30746:Benzoic acid; 17895:Tyrosine; 16349:Citrulline; 15724:Trimethylamine oxide; 45744:N-Acetylneuraminic acid; 17295:Phenylalanine; 15428:Glycine; 36999:Perillic acid; 15741:Succinic acid; 15978:Glycerol phosphate; 16411:Indoleacetic acid
Métabolisme des glycosphingolipides	45744:N-Acetylneuraminic acid
Métabolisme de la glycine, de la sérine et de la thréonine	16857:Threonine; 17115:Serine; 32816:Pyruvic acid; 16015:Glutamic acid; 15428:Glycine; 17724:N,N-Dimethylglycine; 15354:Choline; 30831:Oxobutyric acid; 32398:Glyceric acid; 15811:Phosphoserine; 17750:Betaine; 16467:Arginine; 16919:Creatine; 30915:Ketoglutaric acid; 17549:Aminolevulinic acid; 15729:Ornithine; 16344:Guanidineacetic acid
Métabolisme des glycérophospholipides	16870:Glycerophosphorylcholine; 17553:O-Phosphorylethanolamine; 15354:Choline; 17115:Serine; 15978:Glycerol phosphate
Biosynthèse des glycosphingolipides – série ganglio	45744:N-Acetylneuraminic acid
Glycolyse / Gluconéogenèse	32816:Pyruvic acid; 15978:Glycerol phosphate
Métabolisme du glutathion	16015:Glutamic acid; 17561:Cysteine; 15428:Glycine
Métabolisme des glycérolipides	17553:O-Phosphorylethanolamine; 15978:Glycerol phosphate; 32398:Glyceric acid

Annexe 5 : Résultats des analyses sur l'ensemble des métabolites (annotés et non annotés) pour l'exposition au plomb

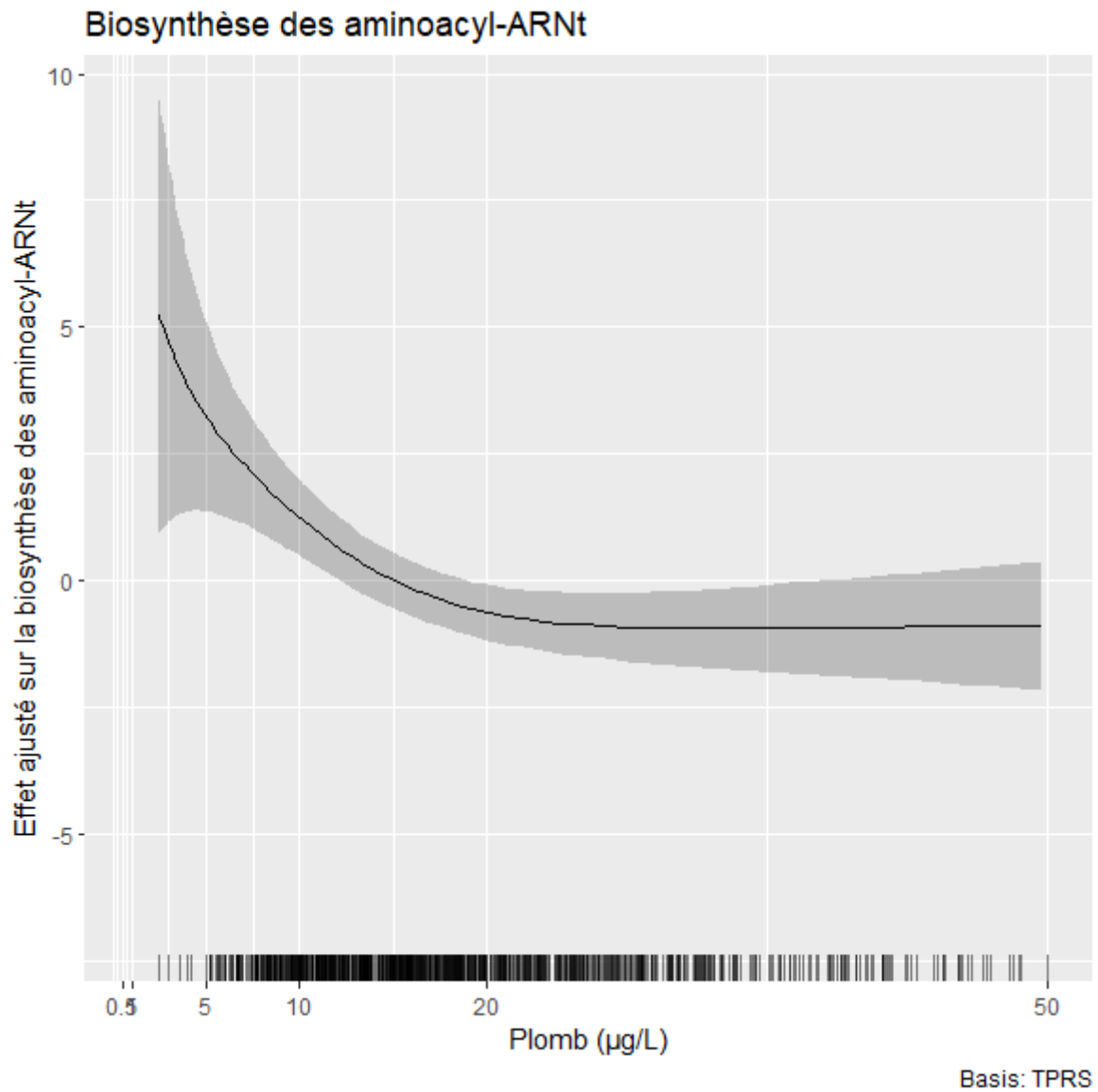


Volcano plot représentant les associations entre l'exposition au plomb et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et lot d'analyse



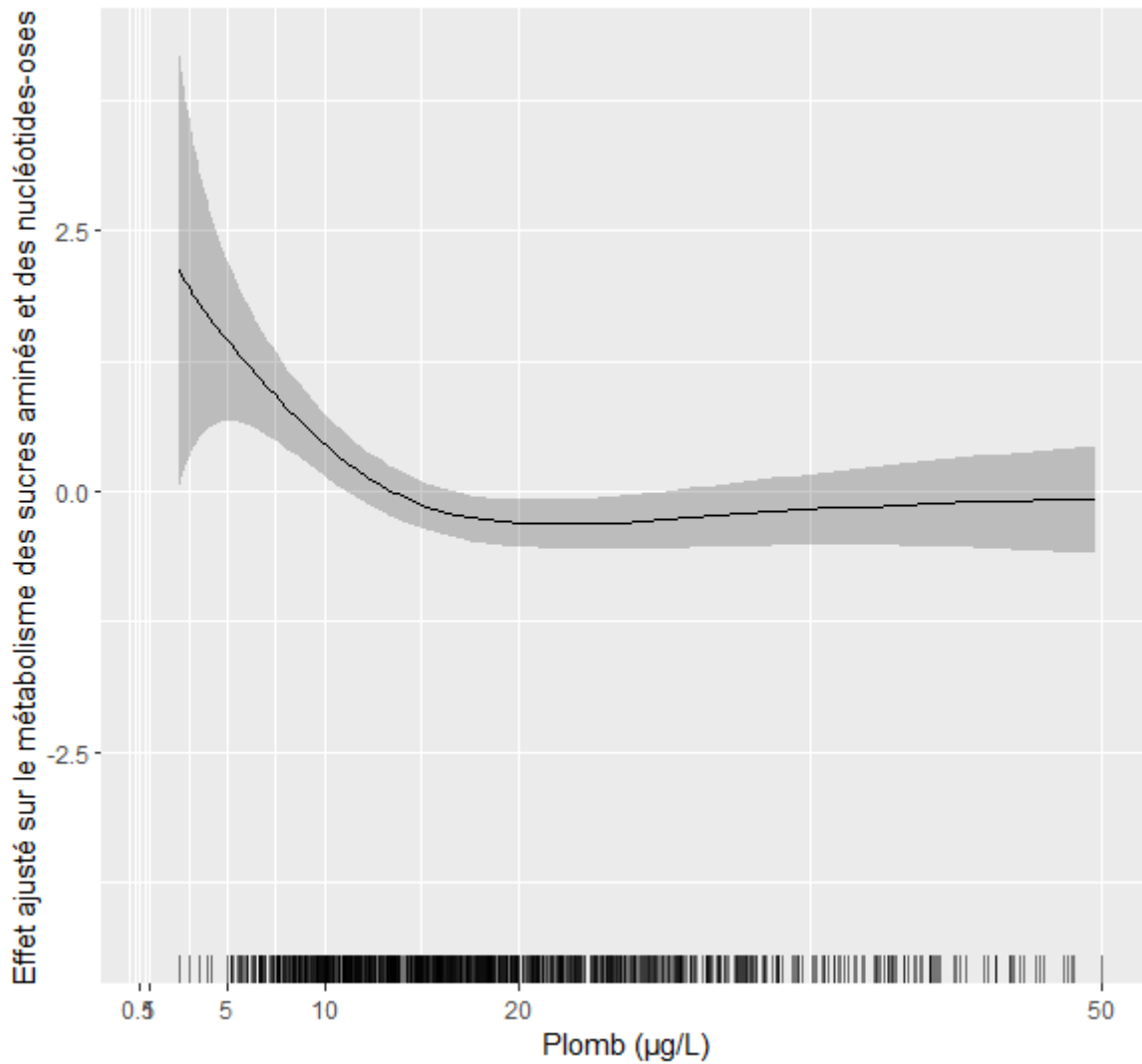
QQ plot représentant les associations entre l'exposition au plomb et les **métabolites**, ajustées sur l'âge, le sexe, la communauté urbaine et lot d'analyse

Annexe 6 : Courbes relation effet-dose pour les autres voies métaboliques (associées au plomb avec la méthode du Z-score et du rapport de vraisemblance)



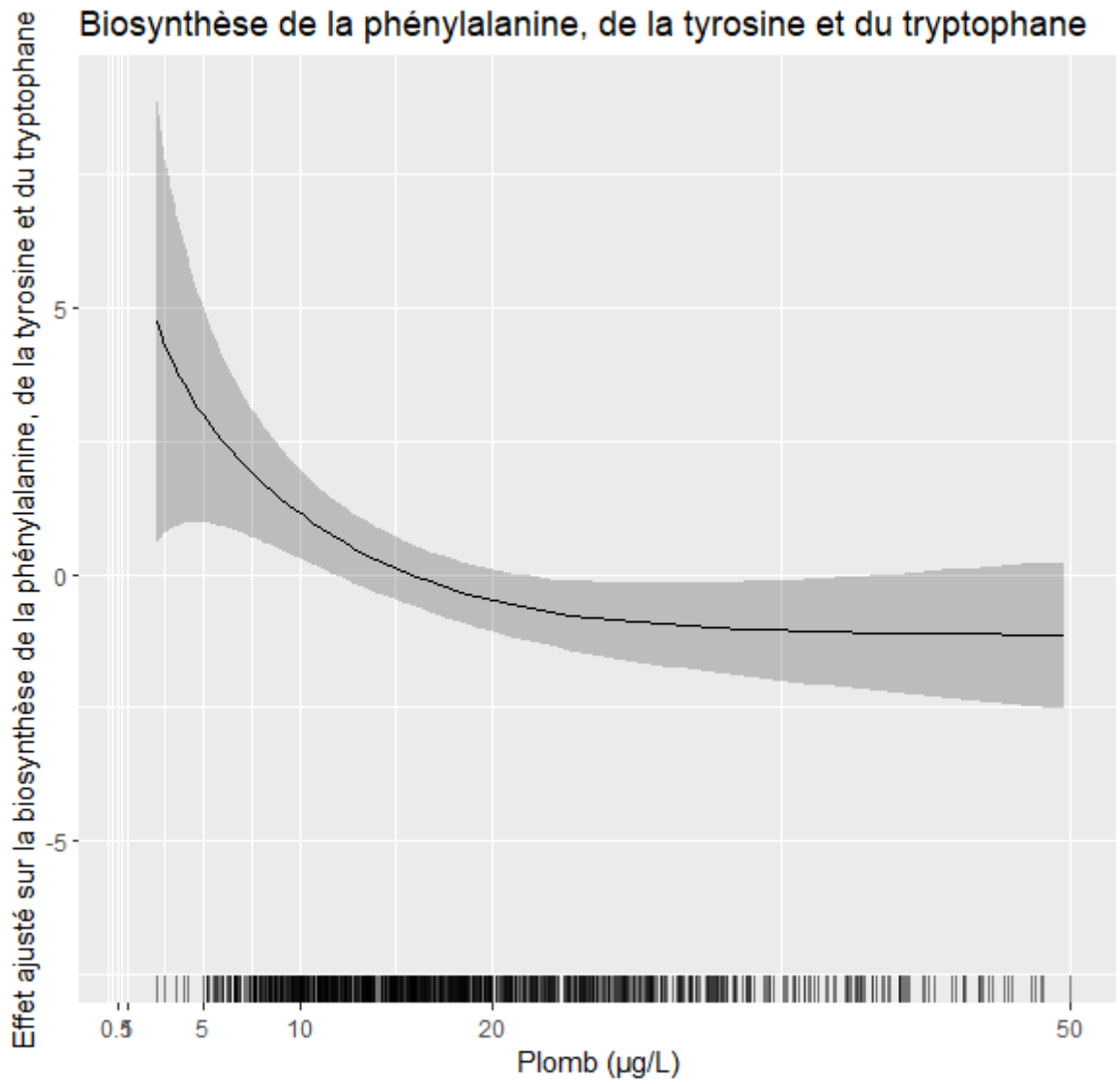
Courbe de la relation effet-dose entre la plombémie et la biosynthèse des aminoacyl-ARNt

Métabolisme des sucres aminés et des nucléotides-oses



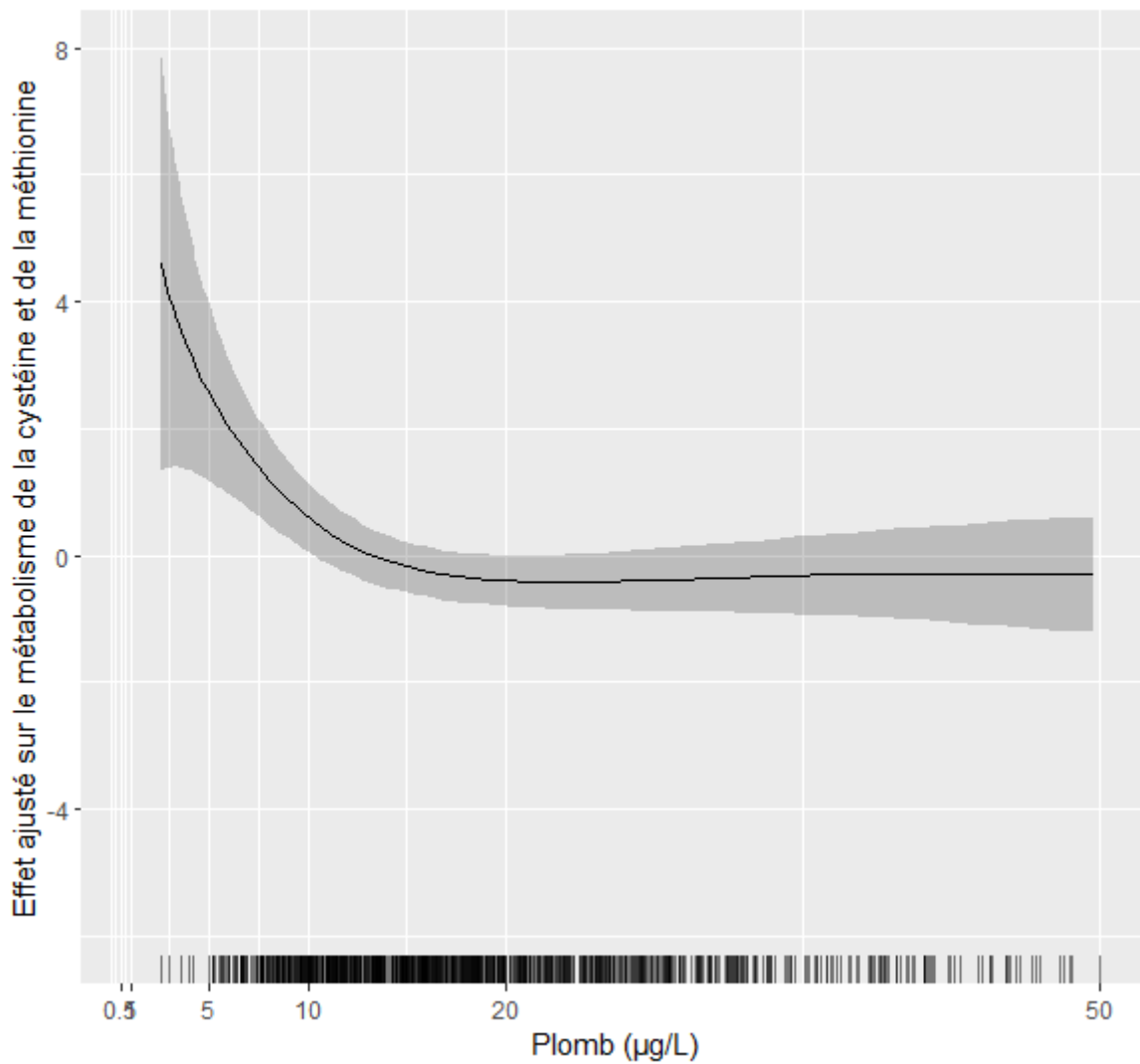
Basis: TPRS

Courbe de la relation effet-dose entre la plombémie et le métabolisme des sucres aminés et des nucléotides-oses

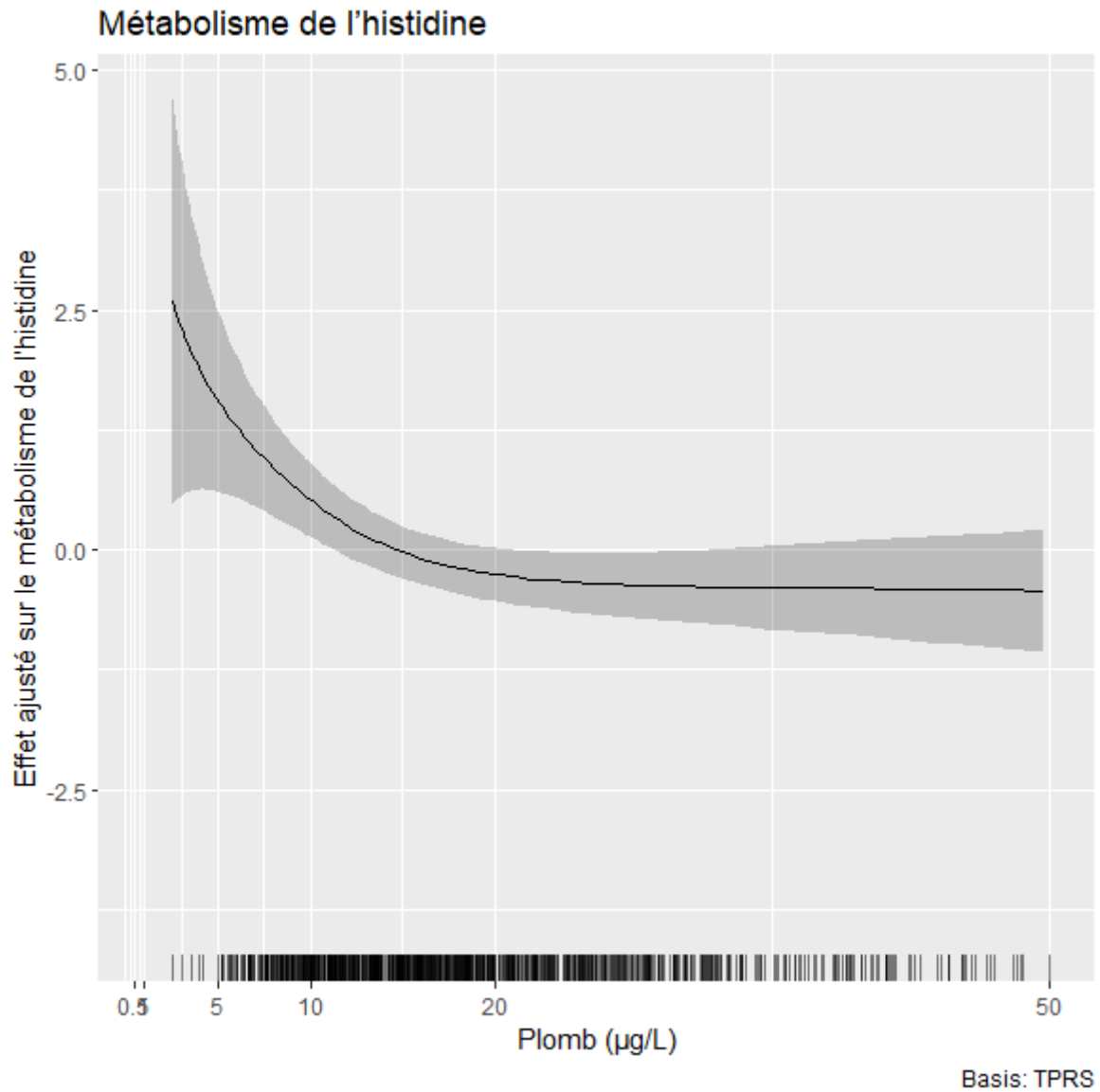


Courbe de la relation effet-dose entre la plombémie et la biosynthèse de la phénylalanine, de la tyrosine et du tryptophane

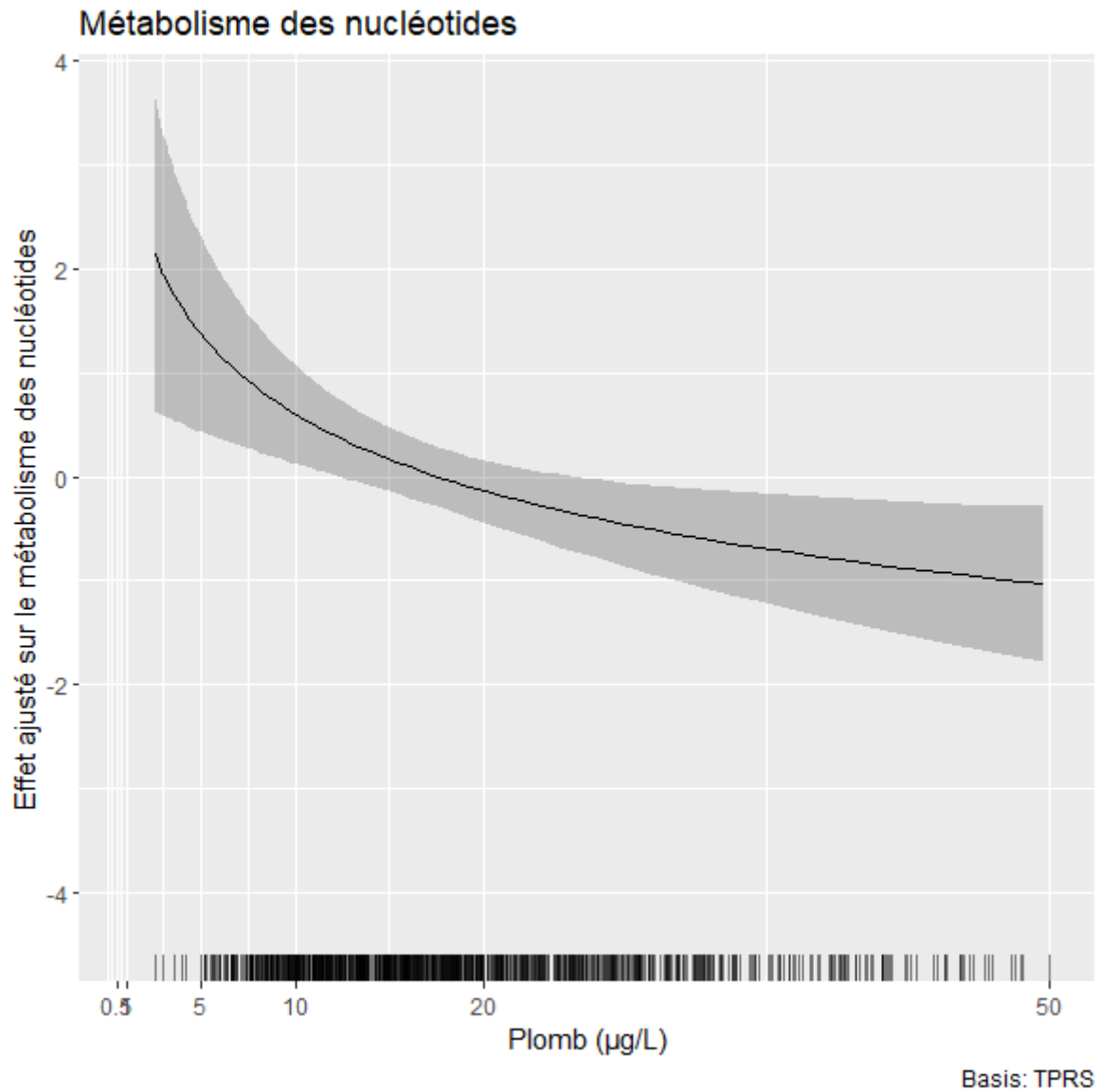
Métabolisme de la cystéine et de la méthionine



Courbe de la relation effet-dose entre la plombémie et le métabolisme de la cystéine et de la méthionine

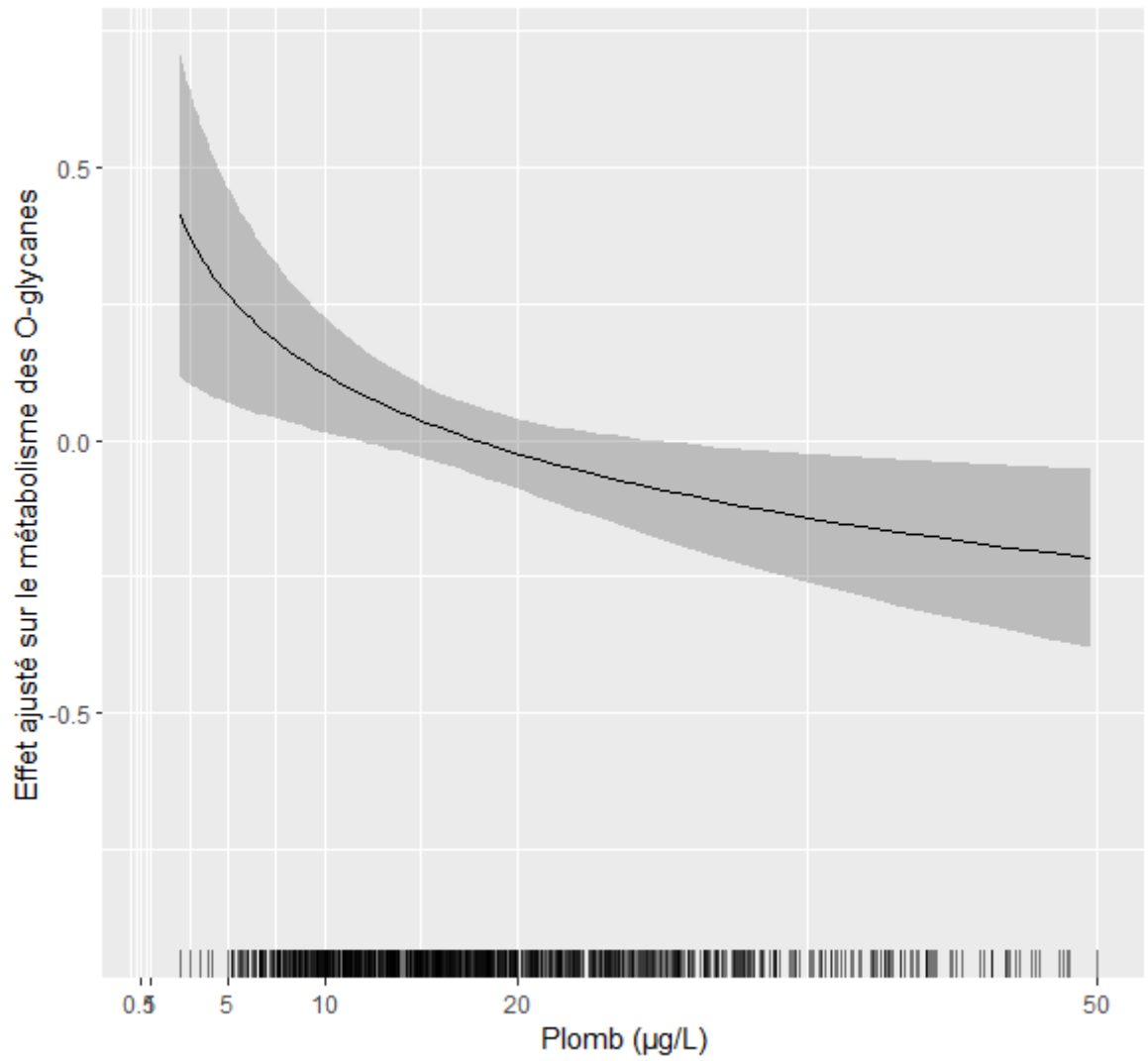


Courbe de la relation effet-dose entre la plombémie et le métabolisme de l'histidine



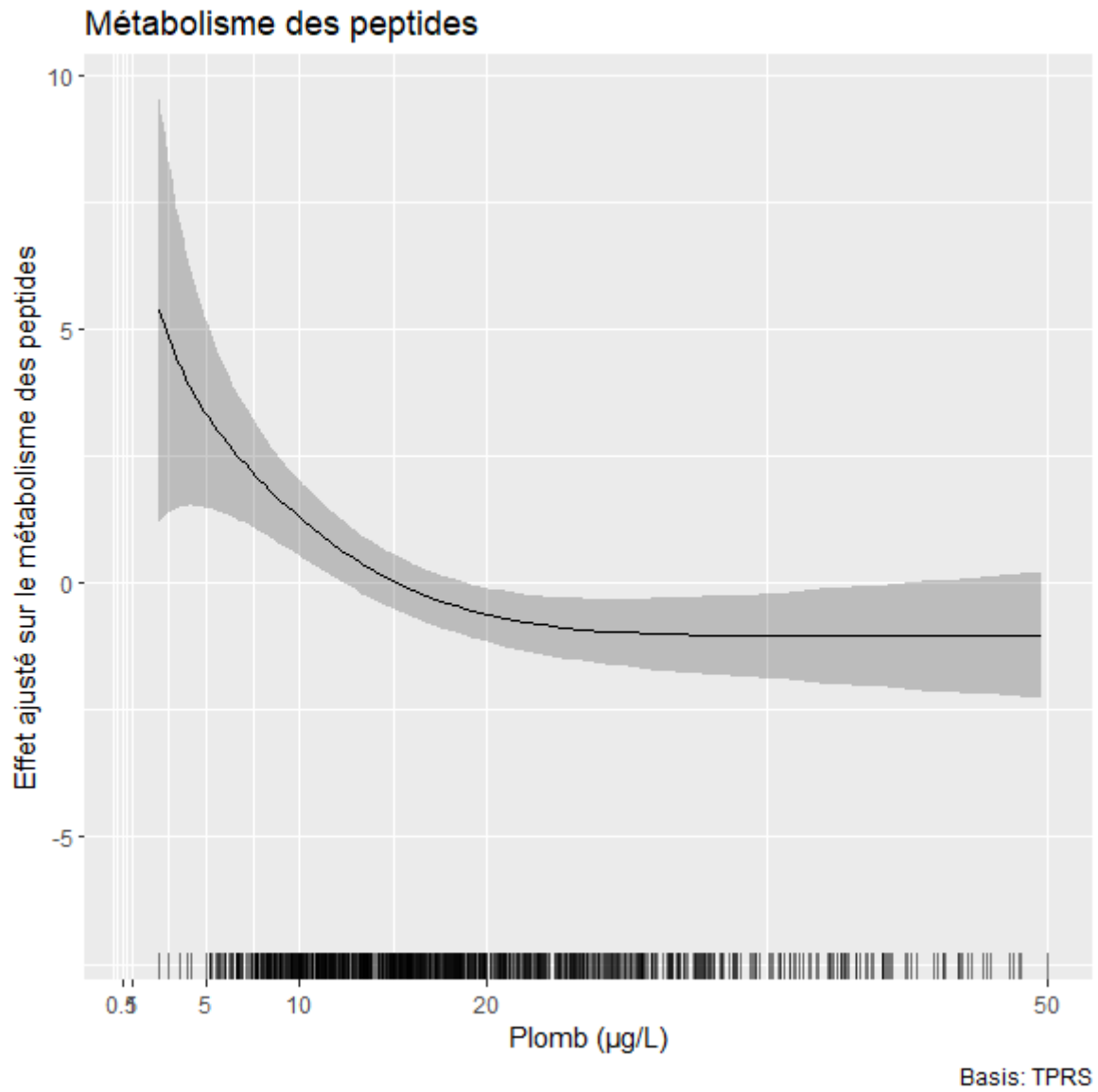
Courbe de la relation effet-dose entre la plombémie et le métabolisme des nucléotides

Métabolisme des O-glycanes



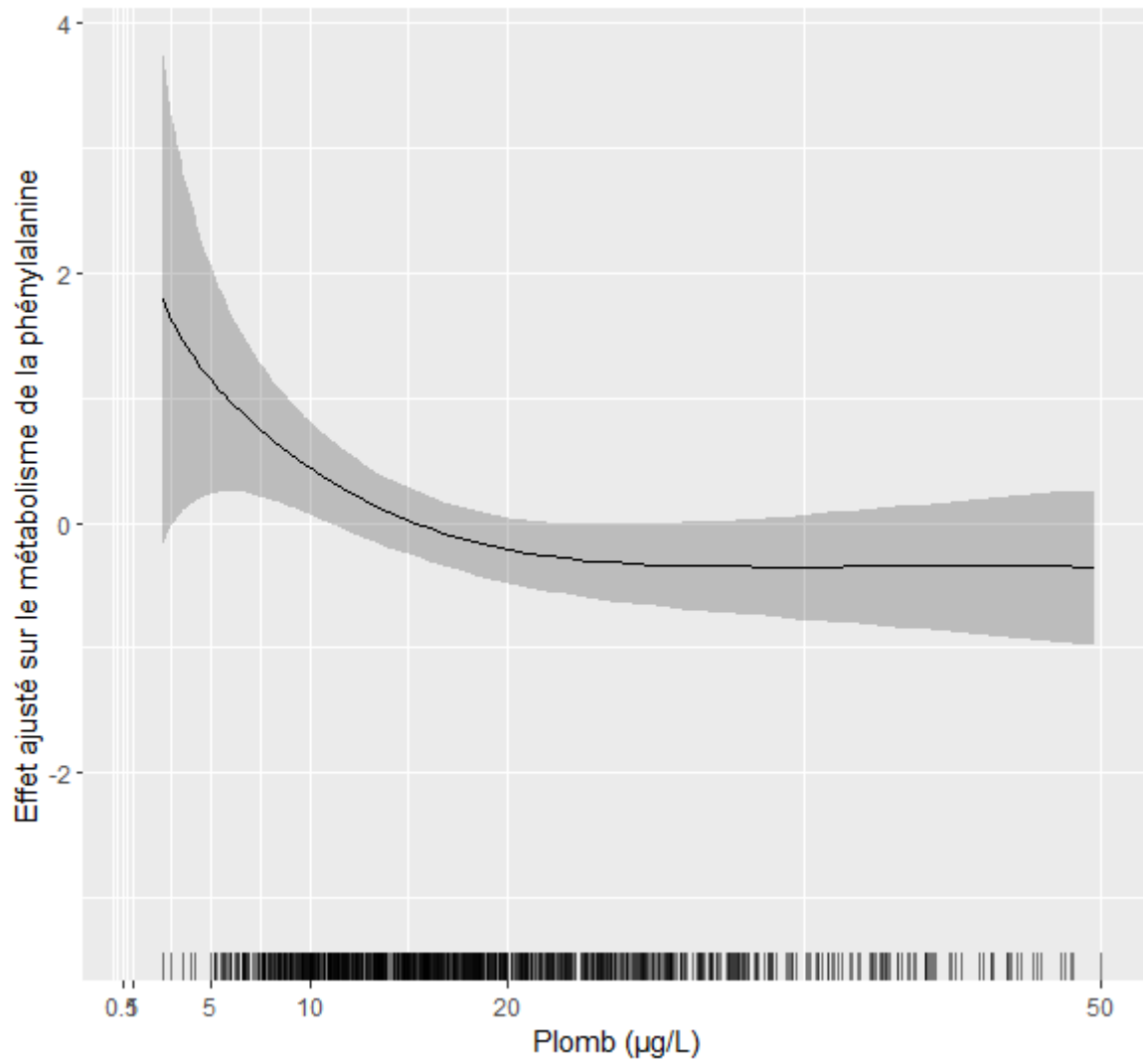
Basis: TPRS

Courbe de la relation effet-dose entre la plombémie et le métabolisme des O-glycanes



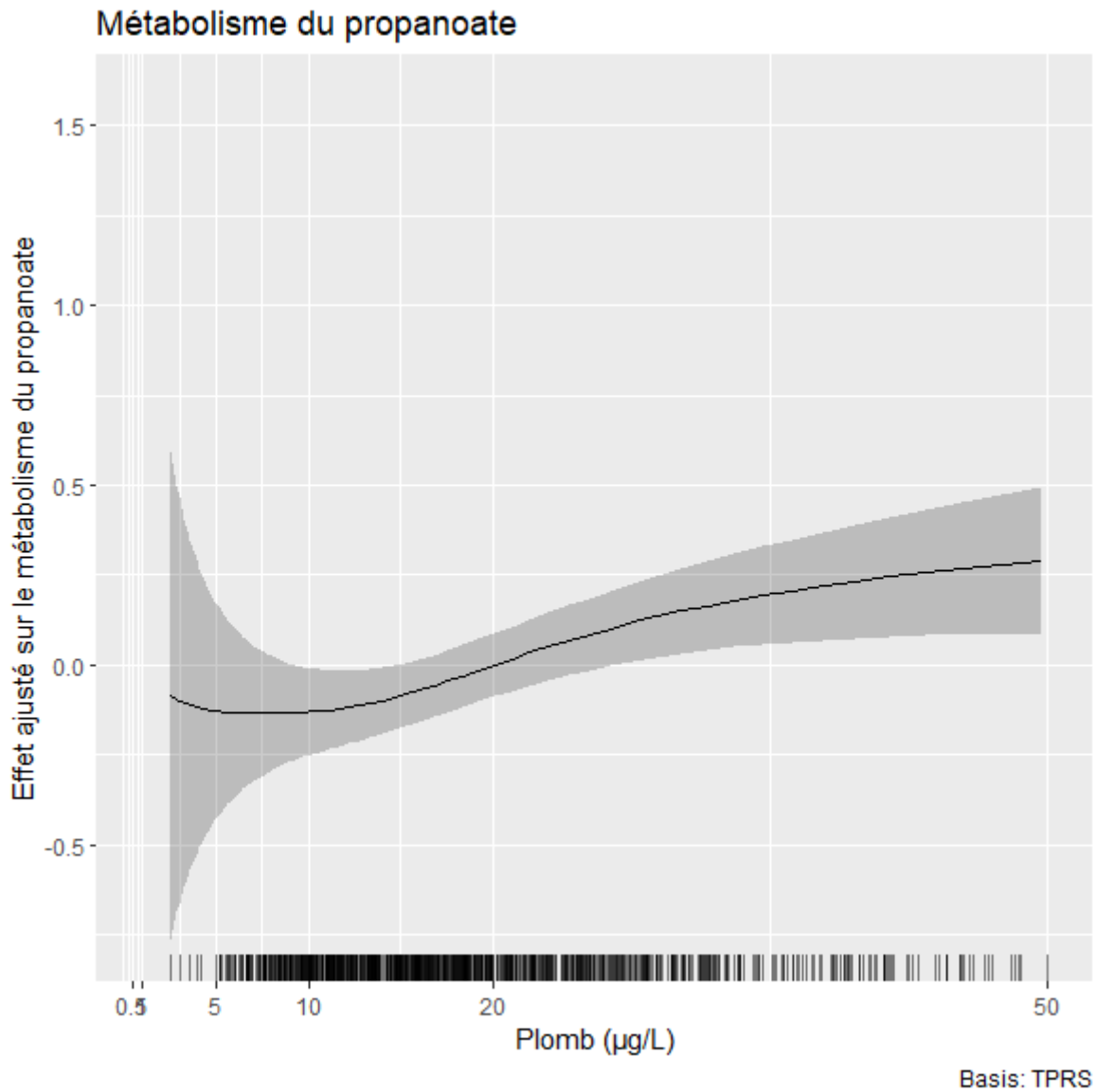
Courbe de la relation effet-dose entre la plombémie et le métabolisme des peptides

Métabolisme de la phénylalanine

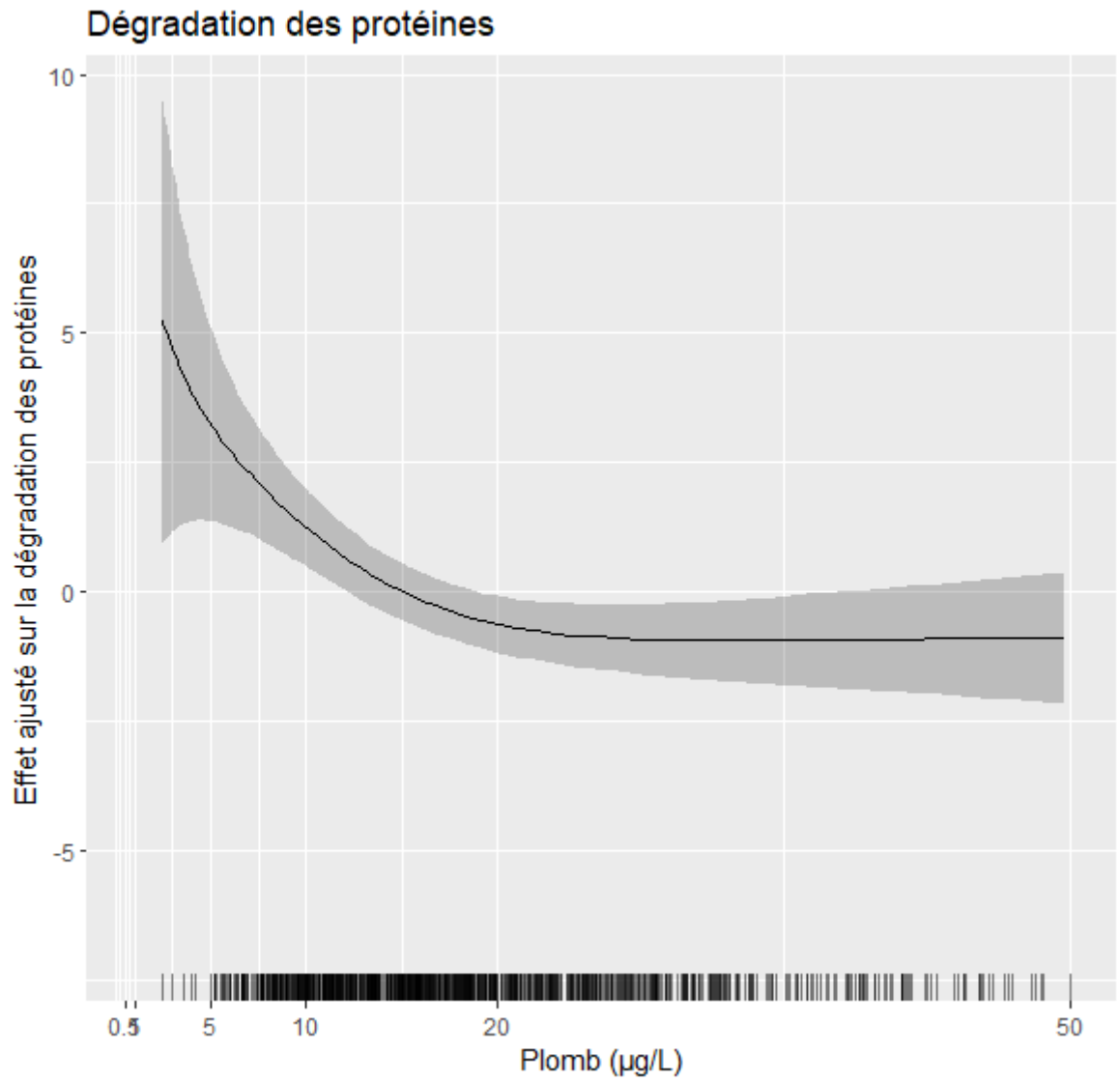


Basis: TPRS

Courbe de la relation effet-dose entre la plombémie et le métabolisme de la phénylalanine



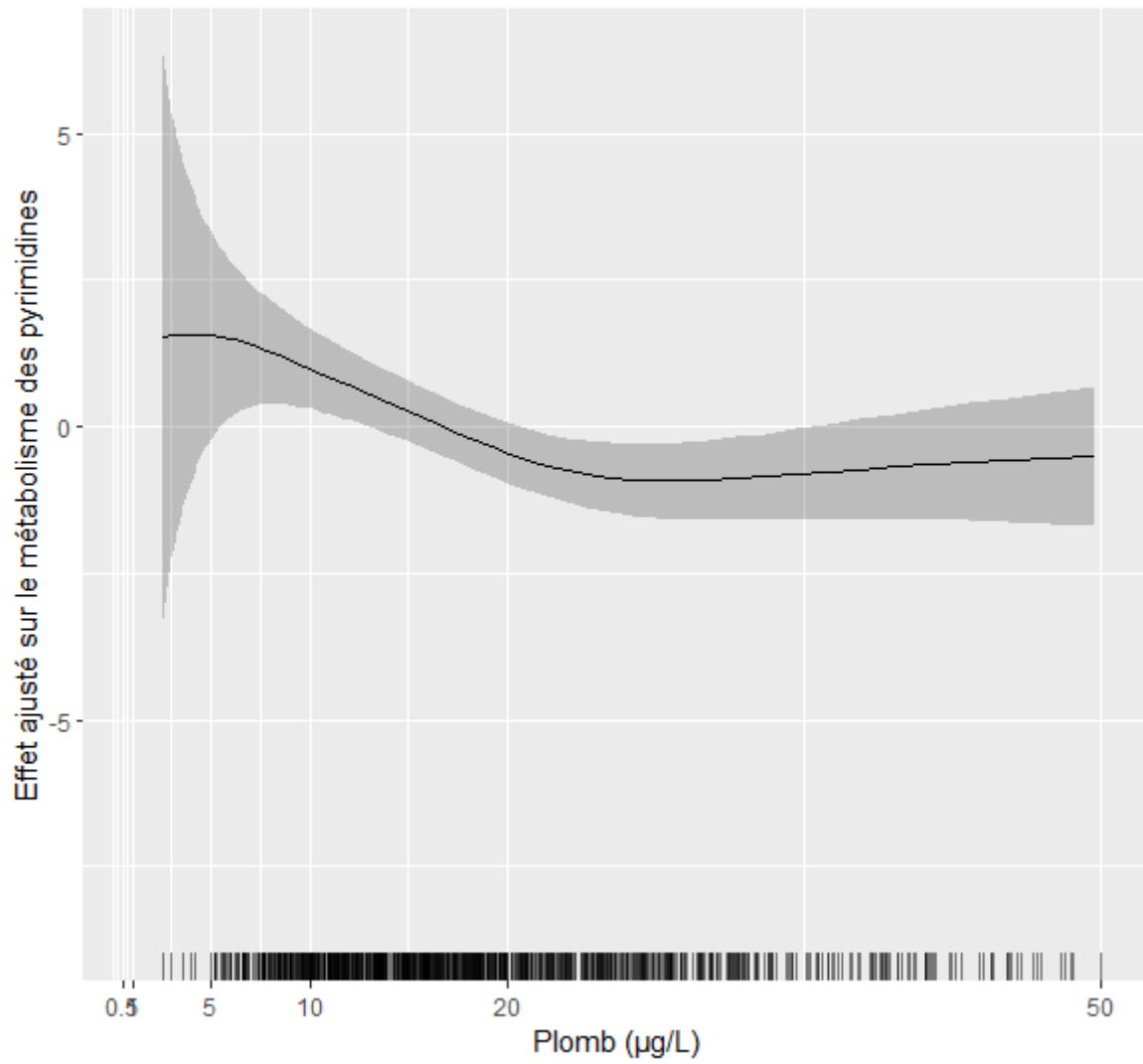
Courbe de la relation effet-dose entre la plombémie et le métabolisme du propanoate



Basis: TPRS

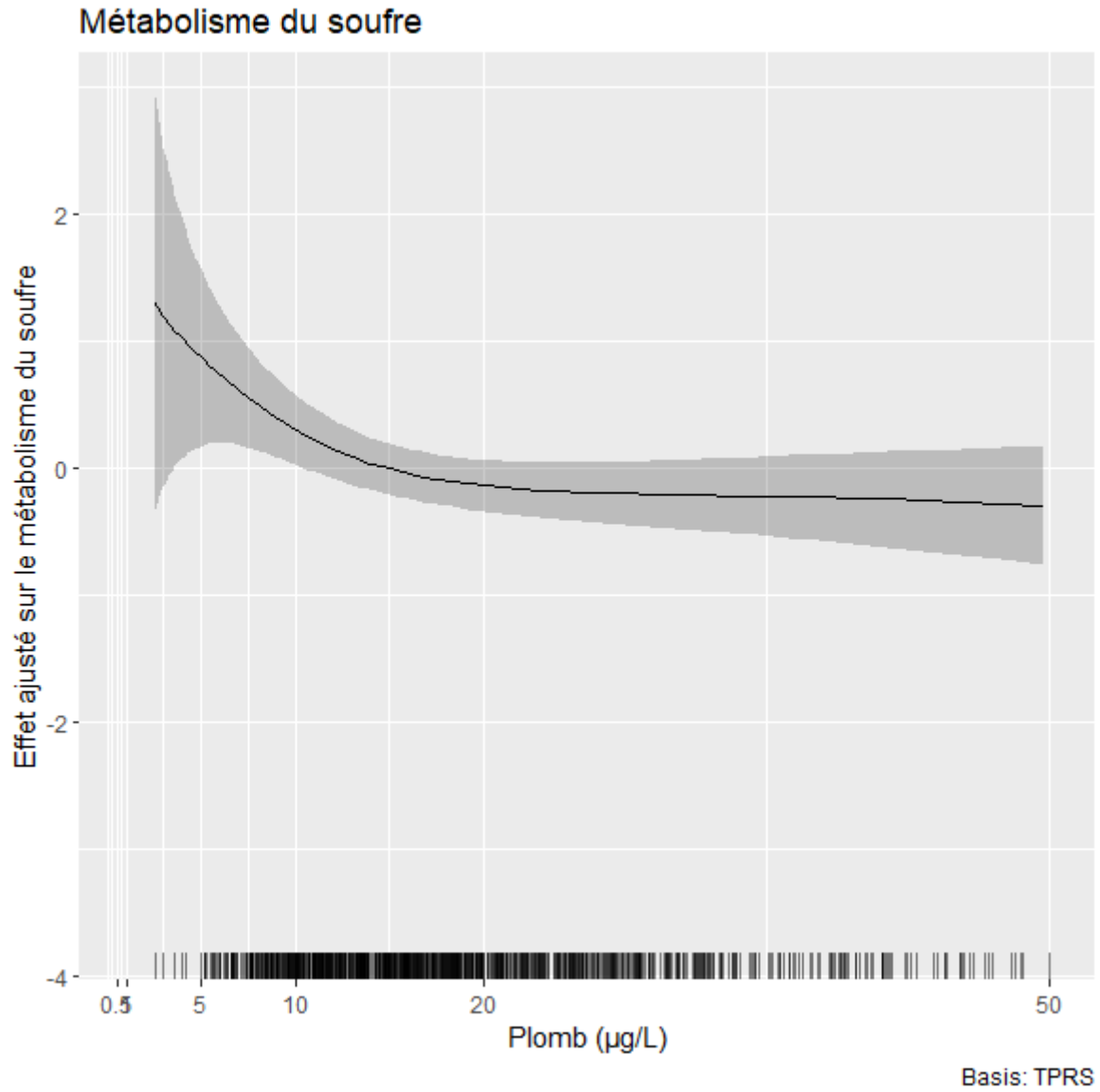
Courbe de la relation effet-dose entre la plombémie et la dégradation des protéines

Métabolisme des pyrimidines



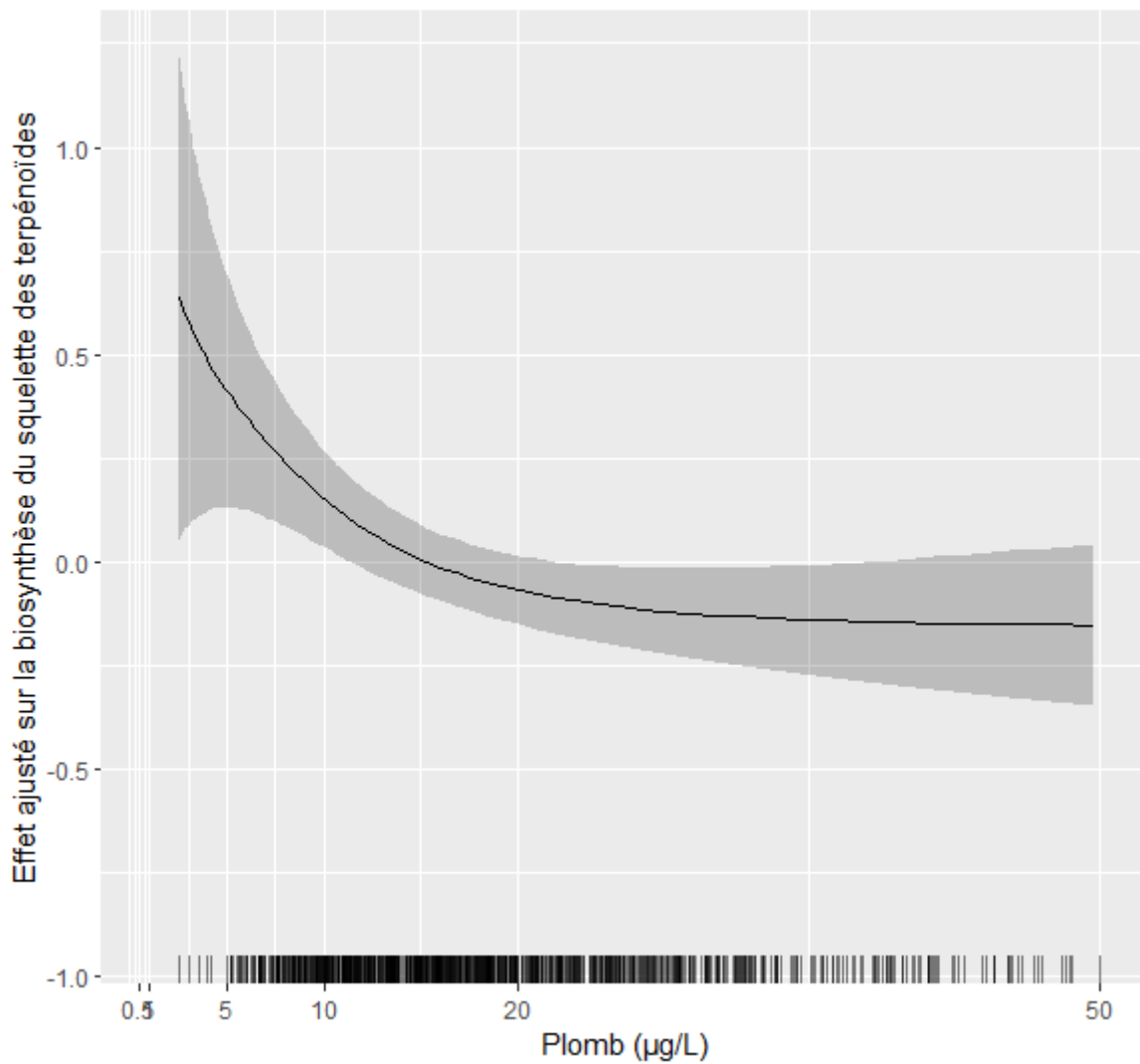
Basis: TPRS

Courbe de la relation effet-dose entre la plombémie et le métabolisme des pyrimidines



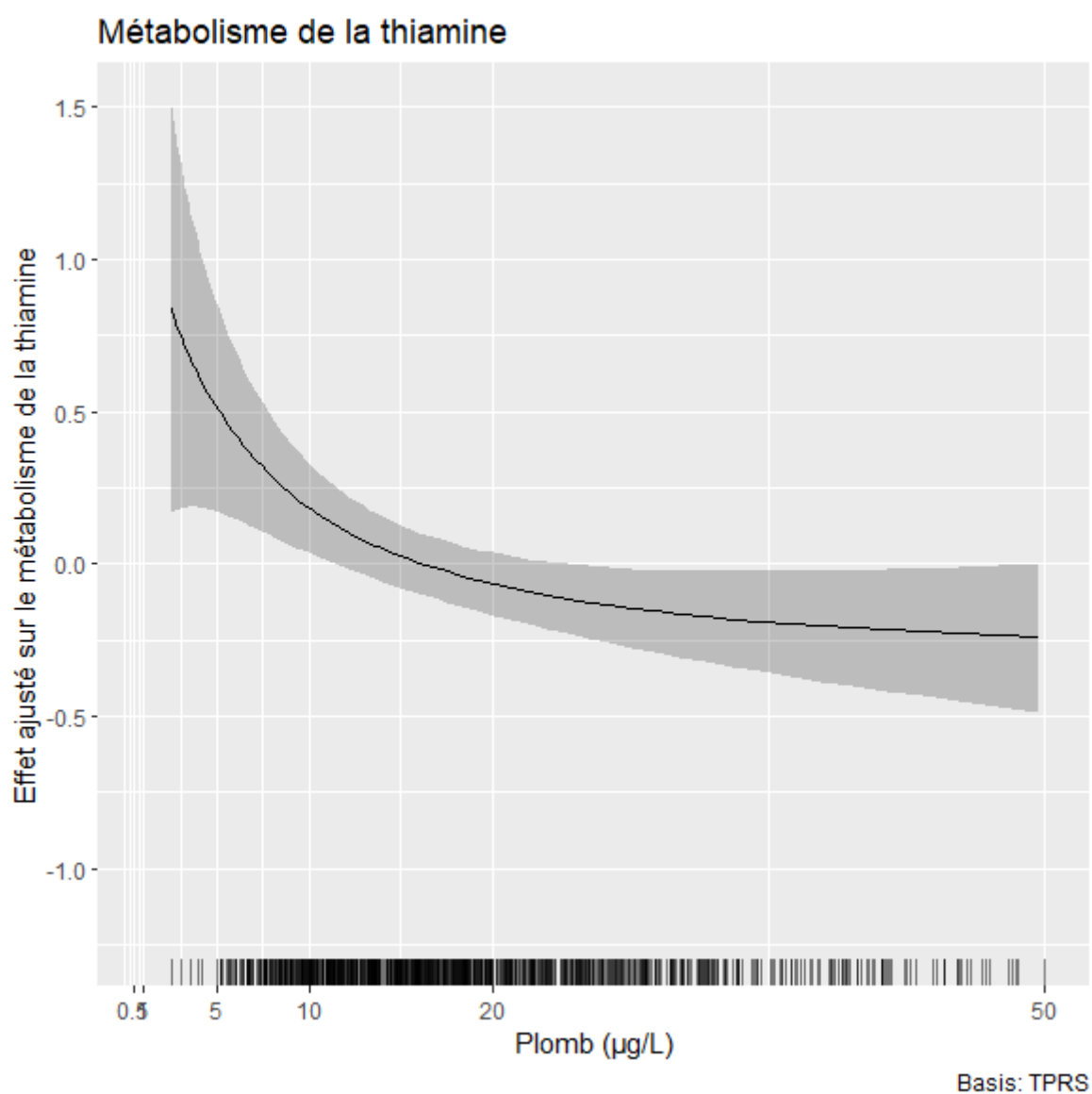
Courbe de la relation effet-dose entre la plombémie et le métabolisme du soufre

Biosynthèse du squelette des terpénoïdes



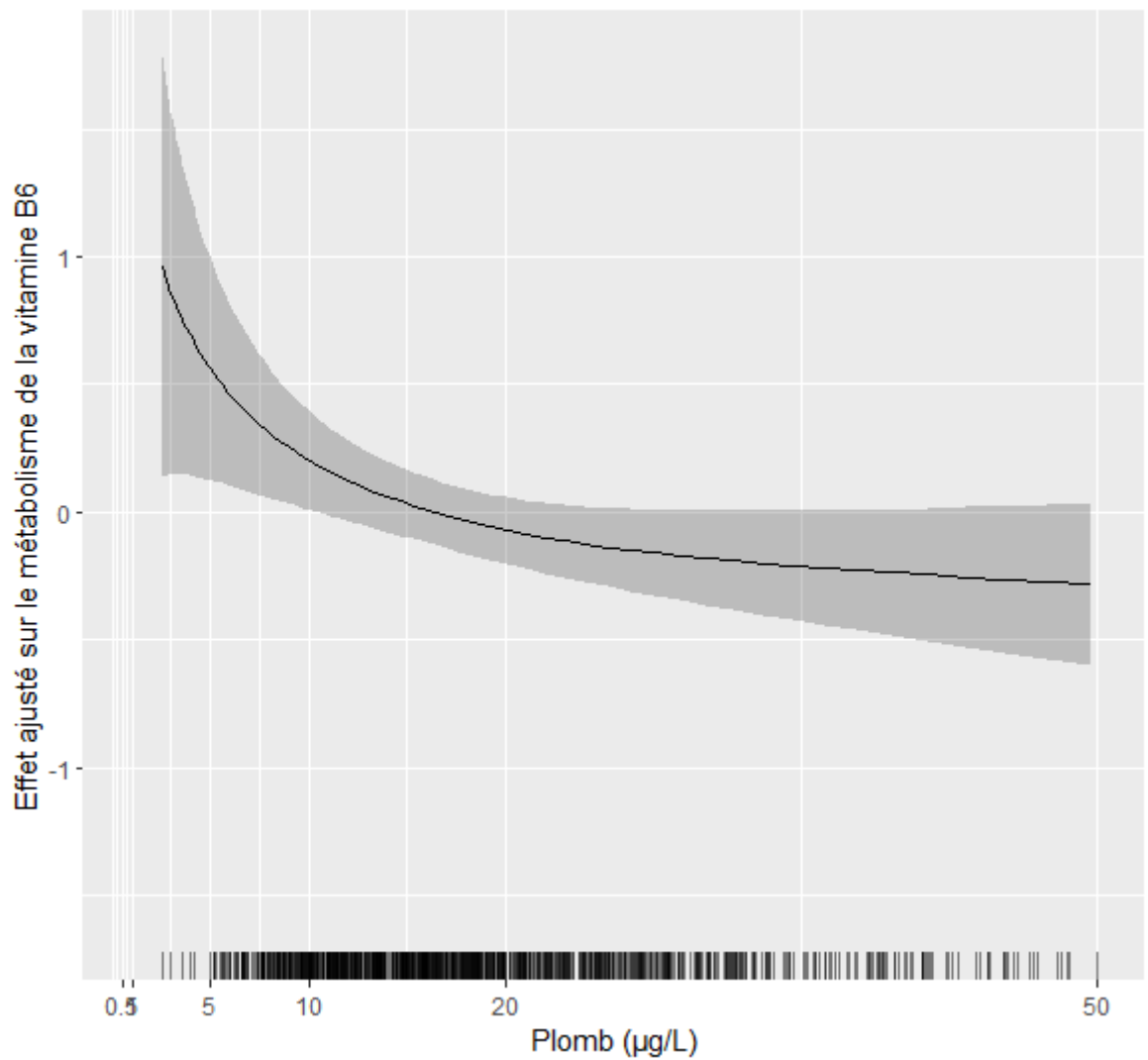
Basis: TPRS

Courbe de la relation effet-dose entre la plombémie et la biosynthèse du squelette des terpénoïdes



Courbe de la relation effet-dose entre la plombémie et le métabolisme de la thiamine

Métabolisme de la vitamine B6



Basis: TPRS

Courbe de la relation effet-dose entre la plombémie et le métabolisme de la vitamine B6