



HAL
open science

Dioxin exposure and breast cancer risk in the E3N cohort : multi-source exposures and timing of exposure

Aurélie Danjou

► **To cite this version:**

Aurélie Danjou. Dioxin exposure and breast cancer risk in the E3N cohort : multi-source exposures and timing of exposure. *Cancer*. Université de Lyon, 2016. English. NNT : 2016LYSE1308 . tel-01662357

HAL Id: tel-01662357

<https://theses.hal.science/tel-01662357>

Submitted on 13 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre NNT : 2016LYSE1308

THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale N° 205
EDISS, École Doctorale Interdisciplinaire Sciences-Santé

Spécialité de doctorat : Épidémiologie, Santé Publique, Recherche sur les services de santé

Discipline : Épidémiologie

Soutenue publiquement le 12/12/2016, par :
Aurélie Marcelle Nicole DANJOU

**Dioxin exposure and breast cancer risk in the E3N cohort:
multi-source exposures and timing of exposure**

Devant le jury composé de :

Pr BALDI Isabelle

PU-PH, Inserm U1219, Université de Bordeaux – Rapporteur

Pr DEGUEN Séverine

PU, HDR, enseignant-chercheur, EHESP, Rennes – Rapporteur

Pr VINEIS Paolo

MD, MRC/PHE Center for Environmental and Health School of Public Health, Imperial College London – Rapporteur

Dr COX David

HDR, CR1, Inserm U1052, CRCL – Examineur

Pr SCHÜZ Joachim

PhD, HDR, Head of Section Environment and Radiation, IARC – Examineur

Pr VIEL Jean-François

MD, PU-PH, HDR, Inserm-IRSET 1085, Université Rennes 1 – Examineur

Pr FERVERS Béatrice

MD, HDR, Pr. associé, Coordinatrice du Département Cancer Environnement, Centre Léon Bérard, Inserm U1052, CRCL, Université de Lyon – Directrice de thèse

Dr DOSSUS Laure

PhD, Section Nutrition and Metabolism, IARC – Co-encadrante

English summary

Breast cancer (BC) is the first malignancy among women. Its incidence has doubled over the past 30 years. Environmental factors with endocrine disruptive properties, such as dioxins emitted from industrial combustion processes, are suspected to affect BC risk. Ingestion of contaminated food and inhalation are the major exposure routes in humans. Epidemiological evidence on the association between dioxin exposure and BC risk remains inconclusive due to methodological limitations. The aim of the thesis was to investigate the association between dioxin exposure and BC risk in the E3N prospective cohort, filling current methodological gaps.

First, we assessed the association between estimated dietary dioxin exposure and BC risk among women from the E3N cohort. Second, we developed a geographic information system (GIS)-based metric to assess airborne dioxin exposure at the individual address level, including proximity to and technical characteristics of industrial sources, exposure duration and prevailing wind frequency. The metric was then applied to each E3N women's addresses from 1990 to 2008, and airborne dioxin exposure was estimated for cases and matched controls from a cohort sub-population (the Rhône-Alpes region). Third, we estimated BC risk associated with cumulative airborne dioxin exposure.

Overall, no statistically significant association was observed, except for a decrease in hormone-independent BC risk. The latter was significant for dietary dioxin exposure. For airborne exposure, we might have lacked statistical power and confirmation at the national level is required. The inverse association with ER-negative BC risk is consistent with experimental evidence.

Keywords: dioxins, breast cancer, dietary exposure, environmental exposure, geographical information system, timing of exposure, epidemiology

French thesis title

**Exposition aux dioxines et risque de cancer du sein
dans la cohorte E3N:
expositions multi-sources et temps d'exposition**

Résumé en Français

Le cancer du sein est le cancer le plus fréquent chez la femme et son incidence a doublé ces 30 dernières années. Les facteurs environnementaux à effet perturbateur endocrinien, tels que les dioxines émises par les activités de combustion industrielle, sont suspectés d'augmenter le risque de cancer du sein. L'alimentation et l'inhalation sont les deux voies majeures d'exposition aux dioxines chez l'Homme. Les données épidémiologiques sur le sujet sont non concluantes, et il existe des limites méthodologiques. Ce travail doctoral avait pour objectif d'étudier l'impact de l'exposition aux dioxines sur le risque de cancer du sein dans la cohorte E3N, en répondant aux limites des études existantes.

Nous avons évalué l'exposition alimentaire aux dioxines puis estimé le risque de cancer du sein associé parmi les femmes de la cohorte E3N. Nous avons ensuite développé un score d'exposition basé sur un système d'information géographique, associant la distance à la source, la durée d'exposition et la fréquence de vent dominant, afin d'évaluer l'exposition environnementale aux dioxines à chaque adresse des femmes entre 1990 et 2008. Le risque de cancer du sein associé au score d'exposition cumulé a été estimé dans une étude cas-témoins nichée dans la cohorte E3N, parmi les femmes ayant résidé en Rhône-Alpes.

Aucune association n'a été observée, à l'exception d'une diminution du risque de cancer du sein hormono-indépendant, retrouvée de façon significative dans l'étude alimentaire. Cette dernière observation est cohérente avec des données expérimentales. Dû à un manque de puissance statistique pour l'exposition aérienne, nos résultats demandent confirmation au niveau national.

Mots-clés : dioxines, cancer du sein, exposition alimentaire, exposition environnementale, système d'information géographique, temps d'exposition, épidémiologie

UNIVERSITE CLAUDE BERNARD LYON 1

Président de l'Université

Vice-président du Conseil d'Administration

Vice-président du Conseil des Etudes et de la Vie
Universitaire

Vice-président du Conseil Scientifique

Directeur Général des Services

M. François-Noël GILLY

M. le Professeur Hamda BEN HADID

M. le Professeur Philippe LALLE

M. le Professeur Germain GILLET

M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Directeur : M. le Professeur J. ETIENNE

Faculté de Médecine et de Maïeutique Lyon Sud – Charles
Mérieux

Directeur : Mme la Professeure C. BURILLON

Faculté d'Odontologie

Directeur : M. le Professeur D. BOURGEOIS

Institut des Sciences Pharmaceutiques et Biologiques

Directeur : Mme la Professeure C.
VINCIGUERRA

Institut des Sciences et Techniques de la Réadaptation

Directeur : M. le Professeur Y. MATILLON

Département de formation et Centre de Recherche en Biologie
Humaine

Directeur : Mme. la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Directeur : M. F. DE MARCHI

Département Biologie

Directeur : M. le Professeur F. FLEURY

Département Chimie Biochimie

Directeur : Mme Caroline FELIX

Département GEP

Directeur : M. Hassan HAMMOURI

Département Informatique

Directeur : M. le Professeur S. AKKOUCHE

Département Mathématiques

Directeur : M. le Professeur G. TOMANOV

Département Mécanique

Directeur : M. le Professeur H. BEN HADID

Département Physique

Directeur : M. Jean-Claude PLENET

UFR Sciences et Techniques des Activités Physiques et Sportives

Directeur : M. Y. VANPOULLE

Observatoire des Sciences de l'Univers de Lyon

Directeur : M. B. GUIDERDONI

Polytech Lyon

Directeur : M. P. FOURNIER

Ecole Supérieure de Chimie Physique Electronique

Directeur : M. G. PIGNAULT

Institut Universitaire de Technologie de Lyon 1

Directeur : M. le Professeur C. VITON

Ecole Supérieure du Professorat et de l'Education

Directeur : M. le Professeur A. MOUGNIOTTE

Institut de Science Financière et d'Assurances

Directeur : M. N. LEBOISNE

Affiliations of the members of the defense thesis jury

Pr Isabelle BALDI

PU-PH

Centre de Recherche Inserm U1219 – Equipe EPICENE

Université de Bordeaux

Dr David COX

CR1 Inserm

U1052, Centre de Recherche en Cancérologie de Lyon

Pr Séverine DEGUEN

PU, HDR, enseignant-chercheur à l'École des Hautes Etudes en Santé Publique, Université Sorbonne Paris Cité et Université Bretagne Loire, Rennes

Dr Laure DOSSUS

PhD, Scientist

Nutrition and Metabolism Section, International Agency for Research on Cancer

Pr Béatrice FERVERS

Professeur associé, MD, PhD, HDR

Coordinatrice du Département Cancer et Environnement, Centre Léon Bérard

Équipe « Signalisation des hormones stéroïdiennes et cancer du sein », UMR Inserm 1052 – CNRS 5286, Centre de Recherche en Cancérologie de Lyon

Pr Joachim SCHÜZ

PhD, HDR

Head of Section Environment and Radiation, International Agency for Research on Cancer

Pr Jean-François Viel

MD, PU-PH, HDR

Institut de Recherche en Santé, Environnement et Travail, Inserm-IRSET 1085, Université Rennes 1

Pr Paolo VINEIS

MD, MPH, FFPH, HDR

Chair in Environmental Epidemiology

MRC/PHE Centre for Environment and Health, School of Public Health, Imperial College London

Author's affiliations and collaborations

The work presented in the thesis was conducted in the following departments:

Département Cancer et Environnement

Centre Léon Bérard

28 rue Laënnec, 69373 Lyon Cedex 08, France

Team “Steroid signalling and breast tumor”, UMR Inserm 1052, CNRS 5286

Cancer Research Centre of Lyon

Université Claude Bernard Lyon 1

28 rue Laënnec, 69373 Lyon Cedex 08, France

The work was achieved with the collaboration of:

Team « Lifestyle, genes and health », Inserm U1018

Center for Research in Epidemiology and Population Health (CESP), Université Paris-Sud

114 rue Edouard-Vaillant, 94805 Villejuif Cedex, France

Team « Epidemiology and Biostatistics », Inserm U1219

Institut de Santé Publique, d'Épidémiologie et de Développement, Université de Bordeaux

146 rue Léo Saignat, 33076 Bordeaux, France

Acknowledgments

I would like to thank my thesis supervisor, Béatrice Fervers, for welcoming me within the Cancer and Environment research department and for having directed this doctoral research. Thank you for kindness, your advice and your guidance.

I would like to express all my gratitude to Laure Dossus, who first gave me the chance to prove myself in this project. Thank you for your kindness, your support, your skills and your trust. Your valuable advice would have helped me until the very end of this work.

I would like to thank Isabelle Baldi, Séverine Deguen and Paolo Vineis for taking the time to review my work. I would also like to thank the other members of the defense thesis jury: David Cox, Joachim Schüz and Jean-François Viel.

Thank you to Barbara Charbotel and David Cox for participating in the thesis follow-up reunions.

Thank you to Béatrice, Laure, Françoise and Thomas for their comments and advice during proofreading of the manuscript.

I would like to address special thank you to my colleagues, with whom I have worked on the GEO3N research project: Camille, Charlotte C, Claire C, Delphine, Elodie F, Guillaume, Hassan, Laure, Lucie, Maxime, Nicolas, Thomas, Tingting and Xavier.

I also would like to thank all my colleagues from the Cancer and Environment research department: Alix, Anne-Sophie, Aude-Marie, Axel, Audrey, Blandine, Cédric B, Charlotte LC, Claire M, Elodie B, Jeffrey, Jennifer, Joane, Julien, Katia, Lidia, Lucie, Marina, Marine, Nora, Nicole, Olivia F, Olivia P, Rémi and Renaud; I apologize for any omission.

I would like to thank all the collaborators on the GEO3N research project, the E3N team in Villejuif: Françoise, Marie-Christine, Camille, Céline, Fabienne, Guy, Maxime and Pascale; Emilie and Karen from ISPED; the École Centrale de Lyon, Air Rhône-Alpes, INERIS, SPF (InVS), DREAL, and INSAValor.

I would like to acknowledge all the financial supports: Université Claude Bernard Lyon 1 and EDISS (for my doctoral fellowship), ADEME, the Ligue contre le Cancer 71, the région Rhône-Alpes and the CLARA.

Thank you to my friends, for their presence and encouragements, from Lyon and Bordeaux.

I would like to address the biggest thank you to my parents and my brothers, for their trust and their support throughout all these years.

I would like to specially thank Cédric, for your affection, your patience and your encouragements.

Remerciements

Je souhaite remercier Béatrice Fervers, de m'avoir accueilli il y a déjà quatre ans au sein du Département Cancer et Environnement, et d'avoir dirigé ce travail. Merci pour votre encadrement et vos conseils.

Je remercie Laure Dossus, qui m'a donné l'opportunité de travailler sur ce projet. Merci pour tes conseils, ta disponibilité, tes compétences et tes encouragements.

Je remercie Isabelle Baldi, Séverine Deguen et Paolo Vineis, d'avoir pris le temps de rapporter ce travail. Je remercie également les autres membres du Jury : David Cox, Joachim Schütz et Jean-François Viel.

Je remercie Barbara Charbotel et David Cox d'avoir participé aux comités de suivi de thèse.

Je remercie Béatrice, Laure, Françoise et Thomas pour leurs remarques et conseils lors de la relecture de ce manuscrit.

Je remercie mes collègues avec qui j'ai travaillé sur le projet GEO3N : Camille, Charlotte C, Claire C, Delphine, Elodie F, Guillaume, Hassan, Laure, Lucie, Maxime, Nicolas, Thomas, Tingting et Xavier.

Je remercie également tous autres mes collègues du Département Cancer et Environnement : Alix, Anne-Sophie, Aude-Marie, Axel, Audrey, Blandine, Cédric, Charlotte LC, Claire M, Elodie B, Jeffrey, Jennifer, Joane, Julien, Katia, Lidia, Marina, Marine, Nora, Nicole, Olivia F, Olivia P, Rémi et Renaud; excusez-moi pour tout oublié.

Je remercie les membres de l'équipe qui dirige la cohorte E3N à Villejuif et avec qui nous collaborons. Merci à Françoise et Marie-Christine ; merci également à Camille, Céline, Fabienne, Guy, Maxime et Pascale.

Je remercie Émilie et Karen de l'ISPED, d'avoir pris le temps de me former aux méthodes statistiques que vous développez dans votre équipe.

Je remercie les collaborateurs au projet GEO3N, ainsi que les membres du comité scientifique : l'École Centrale de Lyon, Air Rhône-Alpes, l'INERIS, SPF (InVS), l'Anses, la DREAL, le CITEPA, le BRGM et INSAValor.

Je remercie l'Université Claude Bernard Lyon 1 et l'EDISS pour m'avoir attribué une bourse de thèse. Je remercie les différents organismes financeurs du projet GEO3N : l'ADEME, la Ligue contre le Cancer 71, la région Rhône-Alpes et le CLARA.

À mes amis, merci pour votre présence et vos encouragements, qu'ils viennent de Lyon ou de Bordeaux.

À mes parents et mes frères, mille mercis pour votre confiance et votre soutien tout au long de ces années.

À Cédric, un merci tout particulier pour ta tendre affection, ta patience et tes encouragements.

Résumé substantiel en Français

Introduction

En France, comme au niveau mondial, le cancer du sein est le cancer le plus fréquent chez la femme et son incidence a augmenté de façon constante ces trente dernières années. Les facteurs de risque du cancer du sein comprennent principalement des facteurs hormonaux et reproductifs ou liés au mode de vie ; les facteurs génétiques expliquent également une faible proportion de cas. Les facteurs environnementaux, présents dans l'air, l'eau ou l'alimentation, sont suspectés d'être impliqués dans l'étiologie du cancer du sein ; d'autant plus que certaines de ces substances agissant comme perturbateurs endocriniens pourraient modifier le risque de cancer hormono-dépendant comme le cancer du sein. Ce travail doctoral s'est plus particulièrement intéressé au rôle dans le développement du cancer du sein des dioxines, dont la TCDD (2,3,7,8-tetrachlorodibenzo-*para*-dioxine) qui a été classée carcinogène certain pour l'Homme par le Centre International de Recherche sur le Cancer (CIRC). Les dioxines sont produites lors de la combustion incomplète de composés chlorés issue d'activités industrielles telles que l'incinération et la métallurgie, ainsi que du chauffage au bois et du trafic routier mais en plus faible proportion. Les dioxines peuvent contaminer les chaînes alimentaires de l'Homme et des animaux d'élevage, et vont s'accumuler dans les tissus adipeux où elles peuvent persister longtemps (plus de 7 ans chez l'Homme). En population générale, l'alimentation est la source principale d'exposition aux dioxines.

Les résultats des études épidémiologiques sur l'association entre exposition aux dioxines et risque de cancer du sein sont contradictoires et présentent certaines limites méthodologiques, telles que le manque d'ajustement sur les facteurs de risque individuels de cancer du sein ; de faibles effectifs ; l'absence de considération de l'exposition alimentaire aux dioxines. À ce jour, aucune étude n'a évalué l'exposition alimentaire aux dioxines en lien avec le risque de cancer du sein. L'évaluation de l'exposition aux dioxines dans les études repose sur des méthodes hétérogènes, qui comprennent l'utilisation de simples métriques comme la distance à la source, la restriction des sources d'exposition aux industries du secteur de l'incinération, l'absence de considération de l'histoire résidentielle et des variations des émissions de dioxines au cours du temps. L'évaluation de l'exposition à des périodes clés du développement mammaire est également à considérer, notamment lors de la gestation, de la puberté et de la grossesse pendant lesquelles la glande mammaire peut être plus vulnérable aux carcinogènes. De même, la période de latence entre exposition et diagnostic doit être assez longue pour le développement tumoral. De nouvelles méthodes d'évaluation des expositions environnementales ont été développées, dont l'utilisation des systèmes d'information géographique (SIG), qui permettent l'évaluation rétrospective des expositions environnementales, au niveau de l'adresse résidentielle des sujets en prenant en compte leur histoire résidentielle, et intégrant des données météorologiques et topographiques.

Objectifs

L'objectif de la thèse était d'étudier l'impact de l'exposition aux dioxines sur le risque de cancer du sein dans la cohorte E3N, en répondant aux limites des études existantes. Les objectifs spécifiques étaient :

1. Évaluer l'exposition alimentaire aux dioxines et estimer le risque de cancer du sein associé à cette exposition dans la cohorte E3N ;
2. Développer et valider un score d'exposition basé sur un SIG pour évaluer l'exposition aérienne aux dioxines ;
3. Estimer le risque de cancer du sein associé à l'exposition aérienne aux dioxines dans une étude cas-témoins nichée dans la cohorte E3N, en se limitant tout d'abord aux participantes ayant exclusivement

résidé en région Rhône-Alpes pendant la période d'étude.

La cohorte E3N

L'étude E3N est une cohorte prospective portant sur près de 100 000 femmes françaises, volontaires, nées entre 1925 et 1950 et étant affiliées à la Mutuelle Générale de l'Éducation Nationale (MGEN). E3N a débuté en 1990 avec pour objectif d'étudier les facteurs de risque de plusieurs maladies chroniques, dont le cancer du sein. Les femmes répondent tous les 2-3 ans à un auto-questionnaire portant sur leur mode de vie et leur état de santé, qui ont régulièrement été mis à jour depuis 1990. À ce jour 11 auto-questionnaires ont été adressés aux participantes. Les femmes ont également complété deux questionnaires d'histoire alimentaire en 1993 et 2005 et fourni des échantillons de sang et de salive. L'histoire résidentielle des participantes a pu être reconstruite à partir des informations contenues dans les questionnaires.

Exposition alimentaire aux dioxines et risque de cancer du sein

(Chapter II, Article #1)

L'exposition alimentaire aux dioxines a été évaluée parmi les 63 830 participantes ayant répondu au questionnaire d'histoire alimentaire en 1993 et ayant été suivies jusqu'en 2008 – pendant cette période, 3 465 cas de cancer du sein ont été diagnostiqués. Les données de consommation alimentaire issues du questionnaire E3N ont été combinées aux données de contamination des aliments par les dioxines fournies par le Conseil Supérieur d'Hygiène Publique de France (CSPHPF), selon une formule recommandée par l'Organisation Mondiale de la Santé (OMS). L'association entre exposition alimentaire aux dioxines et risque de cancer du sein a été estimée par des modèles de Cox, ajustés sur la cohorte de naissance et les facteurs de risque individuels de cancer du sein. Des analyses de sous-groupes ont été réalisées, selon le statut ménopausique et le statut des récepteurs aux estrogènes (ER) et à la progestérone (PR).

À notre connaissance, cette étude est la première à avoir évalué la relation entre exposition alimentaire aux dioxines et risque de cancer du sein. Globalement, aucune augmentation du risque de cancer du sein n'a été observée parmi les femmes de la cohorte E3N. Cependant, une diminution significative du risque de cancer du sein ER-négatif PR-négatif a été observée chez les femmes post-ménopausées.

La limite majeure de notre étude porte sur la non-considération de l'origine des aliments, qui a pu entraîner une sous-estimation de l'exposition alimentaire aux dioxines parmi les femmes résidant à proximité de sources et consommant des aliments produits localement et potentiellement plus contaminés que la moyenne. Cette sous-estimation a pu conduire à un biais de classement non-différentiel et à l'absence d'observation d'une association.

Développement d'un score SIG

(Chapter III, Article #2)

Le développement d'un score SIG pour l'évaluation de l'exposition aérienne aux dioxines des femmes de la cohorte E3N comprenait plusieurs étapes. (1) Un inventaire rétrospectif des sources émettrices de dioxines entre 1990 et 2008 a été effectué, ciblant non seulement les industries du secteur de l'incinération, mais également la production de métaux, la production de chaleur et d'énergie, la production de produits minéraux, la production de produits chimiques et de biens de consommation et les crématoriums. (2) Leurs caractéristiques techniques, telles que la localisation géographique de la cheminée et sa hauteur, les périodes et taux d'activité ou le système de traitement des fumées, ont également été recueillies pendant l'inventaire, permettant leur classement et l'estimation de l'intensité des émissions de dioxines grâce à l'utilisation d'un outil standardisé,

l'UNEP Toolkit. (3) L'histoire résidentielle des participantes et l'adresse des industries ont été géocodées dans le SIG avec ArcGIS, un logiciel de géocodage automatique, qui a montré une bonne précision du géocodage des adresses des participantes de la région Rhône-Alpes.

Afin de définir le score d'exposition aérienne aux dioxines, (4) une revue de la littérature a été effectuée visant à identifier les paramètres pertinents à prendre en compte dans le SIG. Nous avons retenu la distance à la source, la vitesse et la fréquence des vents dominants, la durée d'exposition, la vitesse en sortie d'échappement et la hauteur de cheminée. (5) Chacun de ces paramètres a été intégré dans le SIG selon différentes combinaisons, qui ont été testées par comparaison avec un modèle de dispersion atmosphérique (SIRANE, un modèle de dispersion Gaussien intégrant un module spécifique pour simuler la dispersion des polluants en milieu urbain) sur les villes de Lyon (zone urbaine) et Le Bugey (zone rurale) pour trois années (1996, 2002 et 2008). Des coefficients de Kappa pondérés ont été calculés afin d'estimer la concordance entre les résultats du score SIG considérant différentes combinaisons des paramètres et les concentrations obtenues avec le modèle de dispersion. La meilleure concordance a été observée pour une distance à la source décroissante en $1/d^2$ et une fréquence de vent dominant découpée par angles de 10° et pondérée par les angles adjacents. La vitesse de vent dominant, la vitesse en sortie d'échappement et la hauteur de cheminée n'amélioreraient pas la concordance et n'ont pas été pris en compte dans la définition finale du score SIG.

La limite majeure à l'évaluation de l'exposition aérienne aux dioxines provient d'un inventaire non-exhaustif des sources de dioxines, qui n'a pas considéré le trafic ni les sources mineures ou illégales, dues au brûlage de câbles par exemple, et pourrait entraîner une sous-estimation de l'exposition aérienne aux dioxines. De même, la liste des paramètres inclus dans le SIG et identifiés dans la littérature n'étant pas exhaustive, certains facteurs influençant l'exposition aérienne aux dioxines sont peut-être manquants dans la définition du score, ce qui pourrait biaiser l'estimation de l'exposition aérienne aux dioxines des participantes. Cependant, les résultats du score d'exposition SIG ont montré une concordance « substantielle » à « presque parfaite » avec le modèle de dispersion atmosphérique.

Exposition aérienne aux dioxines et risque de cancer du sein

(Chapter IV, Article #3)

L'association entre l'exposition aérienne aux dioxines et le risque de cancer du sein a été estimée dans une étude cas-témoins nichée dans la cohorte E3N pour laquelle les cas de cancer du sein ont été appariés à des témoins sur l'âge, la date, le statut ménopausique, le département de résidence et l'existence d'un échantillon biologique. Le score SIG a tout d'abord été appliqué aux participantes E3N ayant résidé exclusivement dans la région Rhône-Alpes entre 1990 et 2008, soit 429 cas de cancer du sein appariés à 786 témoins. L'estimation du risque de cancer du sein a été effectuée par des modèles de régression logistique conditionnelle, ajustés sur les facteurs de risque de cancer du sein. Des analyses de sous-groupes ont été réalisées selon le statut ménopausique, le statut des récepteurs ER et PR, l'indice de masse corporelle (IMC), l'âge à la première grossesse, l'âge de ménarche et l'allaitement. Les modèles ont été ajustés sur l'exposition alimentaire aux dioxines. Le risque de cancer du sein a également été estimé en fonction du temps avant le diagnostic par la modélisation d'une fonction de poids.

Aucune association n'a été observée entre l'exposition aérienne aux dioxines et le risque de cancer du sein, de même pour l'estimation du risque en fonction du temps avant le diagnostic. Une diminution du risque de cancer du sein ER-négatif statistiquement non-significative a été observée. L'ajustement sur l'exposition alimentaire aux dioxines ne modifiait l'estimation du risque. Une diminution du risque de cancer du sein a été observée parmi les femmes de faible IMC, jeunes à la première grossesse et ayant allaité. Cependant l'interprétation de nos résultats est difficile du fait du

manque de puissance statistique ; ces résultats doivent être confirmés à plus grande échelle, en particulier dans une étude sur l'ensemble des cas de cancer du sein E3N.

Les limites citées plus haut sur la définition du score SIG ont possiblement entraîné une sous-estimation de l'exposition aérienne aux dioxines qui a pu conduire à un biais de classement non-différentiel et à l'absence d'observation d'une association. Les résultats de l'estimation du risque de cancer du sein en fonction du temps avant le diagnostic sont préliminaires et requièrent à cette date d'autres analyses pour conclure.

Exposition aux dioxines et risque de cancer du sein dans la cohorte E3N

Dans l'ensemble, aucune association n'a été observée entre le risque de cancer du sein et l'exposition alimentaire ou aérienne aux dioxines. Une diminution du risque de cancer du sein hormono-indépendant a été retrouvée associée aux deux voies d'exposition, et de façon statistiquement significative pour l'exposition alimentaire aux dioxines. Nos résultats sont en accord avec des études ayant démontré l'action antiproliférative de la TCDD sur des cellules cancéreuses *in vitro* ou sur des glandes mammaires animales. Une diminution du risque de cancer du sein hormono-indépendant a également été observée dans des études épidémiologiques portant sur l'exposition à certains composés organochlorés et perturbateurs endocriniens (pesticides et polychlorobiphényles (PCBs)). Une étude cas-témoins a également mis en évidence une diminution statistiquement significative du risque de cancer du sein parmi des femmes de plus de 60 ans et vivant à proximité d'une unité d'incinération d'ordures ménagères ; cependant aucune stratification n'a été effectuée sur le statut des récepteurs aux estrogènes et à la progestérone des tumeurs.

Conclusion

Ce travail doctoral avait pour objectif d'étudier l'association entre exposition aux dioxines et risque de cancer du sein dans la population générale, en répondant aux limites méthodologiques de la littérature existante et d'améliorer les connaissances à travers une recherche interdisciplinaire. À notre connaissance, notre étude était la première à estimer l'association entre exposition alimentaire aux dioxines et risque de cancer du sein dans la population générale. Le travail méthodologique concernant l'évaluation de l'exposition aérienne aux dioxines a permis le développement d'un outil standardisé pour l'évaluation de l'exposition environnementale aux dioxines, qui pourra également être adapté à l'évaluation d'autres polluants en relation avec le cancer du sein ou d'autres pathologies.

Table of contents

CHAPTER I. INTRODUCTION.....	24
1. Scientific background.....	25
1.1. Breast cancer among women.....	25
1.1.1. Tumor pathology.....	25
1.1.2. Descriptive epidemiology	25
1.1.3. Etiology.....	28
1.2. Environmental exposure to air pollutants with endocrine disruptive properties	34
1.2.1. Air pollutants with endocrine disruptive properties	34
1.2.2. Dioxins	35
1.3. Dioxin exposure and breast cancer risk.....	37
1.4. Limitations and issues of existing studies on dioxin exposure and breast cancer	39
2. Objectives of the thesis	40
2.1. Objectives.....	40
2.2. Thesis context.....	41
2.2.1. The GEO3N project	42
2.2.2. The E3N cohort study	44
CHAPTER II. DIETARY DIOXIN EXPOSURE AND BREAST CANCER RISK.....	48
1. Introduction.....	49
2. Article #1.....	50
3. Conclusion.....	66
CHAPTER III. CHARACTERIZATION OF AIRBORNE DIOXIN EXPOSURE	68
1. Introduction.....	69
2. Methods.....	70
2.1. Inventory and characterization of dioxin emitting sources	70
2.1.1. Inventory of dioxin emitting sources	70
2.1.2. Characterization of dioxin emitting sources.....	70
2.2. Geocoding	71
2.2.1. Geocoding of the participants' residential history	71
2.2.2. Geolocation of the dioxin emitting sources.....	71
2.3. GIS-based airborne dioxin exposure assessment	72
2.3.1. Literature review	72
2.3.2. Modeling of airborne dioxin exposure.....	72

2.3.3. Selection of parameters' combination to be included in the GIS-based metric	73
3. Results	73
3.1. Inventory and characterization of dioxin emitting sources	73
3.2. Geocoding	74
3.2.1. Geocoding of the participants' residential history	74
3.2.2. Geolocation of the dioxin emitting sources.....	74
3.3. GIS-based airborne dioxin exposure assessment	74
3.3.1. Literature review	74
3.3.2. GIS parameters.....	75
3.3.3. Parameters' combination to be included in the GIS-based metric	76
4. Article #2.....	77
5. Discussion	105
CHAPTER IV. AIRBORNE DIOXIN EXPOSURE INDEX AND BREAST CANCER RISK. 108	
1. Introduction	109
2. Article #3.....	110
3. Additional analysis.....	141
3.1. Time-dependent risk estimation	141
3.2. Questionnaire on Lifetime Residential History.....	141
3.2.1. Description of the QRH	141
3.2.2. Analyses of the QRH	142
3.2.3. Geocoding of the lifetime residential and occupational addresses.....	142
4. Conclusion.....	142
CHAPTER V. DISCUSSION	145
1. Discussion	146
2. Research perspectives	149
2.1. The GEO3N research project	149
2.2. The XENAIR research project	150
3. Public Health impact	151
4. Conclusion.....	152
References	154
Appendices	168
Appendice 1. Literature review: list of algorithms.....	168
Appendice 2. Literature review: presentation of 7 articles selected.....	170

Appendice 3. GIS-based metric tests: weighted kappa coefficients and 95% CI.....	172
Appendice 4. Questionnaire on lifetime residential history	177
Appendice 5. Results of the analyses of the QRH in the Rhône-Alpes region	190
Appendice 6. Publication by Binachon et al. 2014	196

List of abbreviations

ADEME, French Environment and Energy Management Agency
AhR, aryl hydrocarbon receptor
ANSES, French Agency for Food, Environmental and Occupational Health and Safety
Arnt, AhR nuclear translocator
B[a]P, benzo[a]pyrene
BMI, body mass index
BPAC, Breast Cancer Association Consortium
BPC3, Breast and Prostate Cancer Cohort Consortium
BRGM, French geographical survey
CESP, Center for Research in Epidemiology and Population Health
CGHFBC, the Collaborative Group on Hormonal Factors in Breast Cancer
CI, confidence interval
CITEPA, interprofessional technical center for studies on air pollution
CLARA, Cancéropôle Lyon Auvergne Rhône-Alpes
CNIL, French national commission for computed data and individual freedom
CRCL, Cancer Research Centre of Lyon
CSHPF, French High Council on Public Health
DDT, dichlorodiphenyltrichloroethane
DES, diethylstilbestrol
DGAL, Directorate for food of the French Agriculture ministry
DGCCRF, General Directorate for Competition Policy, Consumer Affairs and Fraud Control
DHQ, diet history questionnaire
DRE, dioxin response element
DREAL RA, Direction régionale de l'environnement, de l'aménagement et du logement Auvergne-Rhône-Alpes
E3N, Etude Epidémiologique après de femmes de l'Education Nationale
EDC, endocrine-disrupting compound
EPA, Environmental Protection Agency
EPIC, European Prospective Investigation into Cancer and Nutrition
ER, estrogen receptor
ERE, estrogen response element
ESRI, Environmental System Research Institute
GIS, geographical information system
GWAS, genome-wide association studies
HR, hazard ratio
IARC, International Agency for Research on Cancer
IGN, National Geographic Information Institute
IOM, Institute of Medicine

INED, National research Institute of Demography and Population
INERIS, National competence center for industrial safety and environmental protection
INSEE, National Institute of Statistics and Economic Studies
Inserm, Institut national de la santé et de la recherche médicale
ISPED, Institut de Santé Publique, d'Épidémiologie et de Développement (Bordeaux School of Public Health)
LIBCSP, Long Island Breast Cancer Study Project
LUR, land use regression
MGEN, Mutuelle Générale de l'Éducation Nationale
MHT, menopausal hormone therapy
MSWI, municipal solid waste incinerator
NHL, non-Hodgkin lymphoma
OCDD, octachlorodibenzo-*para*-dioxin
OR, odds ratio
ORS, Observatoire Régional de la Santé
PAH, polycyclic aromatic hydrocarbon
PCB, polychlorinated biphenyl
PCDD, polychlorodibenzo-*para*-dioxin
PCDF, polychlorodibenzofuran
POP, persistent organic pollutant
PR, progesterone receptor
QRH, questionnaire on residential history
ROS, reactive oxygen species
RR, relative risk
SD, standard deviation
SMR, standardized mortality ratio
SNP, single nucleotide polymorphism
TCDD, 2,3,7,8-tetrachlorodibenzo-*para*-dioxin
TDS, total diet study
TEF, toxic equivalence factor
TEQ, toxic equivalence quotient
UNEP, United Nations Environment Program
WHO, World Health Organization
WK, weighted kappa coefficient
XRE, xenobiotic response element

List of figures

Figure 1. Estimated age-standardized female breast cancer incidence rates (per 100,000) worldwide in 2012. GLOBOCAN 2012, International Agency for Research on Cancer.....	26
Figure 2. Estimated age-standardized female breast cancer mortality rates (per 100,000) worldwide in 2012. GLOBOCAN 2012, International Agency for Research on Cancer.....	26
Figure 3. Evolution of the number of incident cases and breast cancer-related deaths between 1980 and 2015 (prediction). Evolution of age-standardized incidence and mortality rates (per 100,000) between 1980 and 2015. Institut National du Cancer, 2015	27
Figure 4. Number of new breast cancer cases and breast cancer-related deaths among women in France by age group – prediction for 2015. Institut National du Cancer 2015	27
Figure 5. Dioxin-estrogen pathway crosstalk. Adapted from Guo 2009.....	36
Figure 6. Calendar of the E3N self-administered questionnaires, 1990-2014 (last update on 07-2016).	44
Figure 7. Extracts from the diet history questionnaire sent in 1993 (quantitative section, page 6) and the photo-booklet (page 24). E3N study, France.....	45
Figure 8. Extract from the diet history questionnaire sent in 1993 (qualitative section, page 17). E3N study, France	46
Figure 9. Evolution of the sum of annual dioxin emission intensities estimated between 1990 and 2008, per activity sectors, among the source-periods inventoried within the Rhône-Alpes region. GEO3N project. France, 1990-2008.....	74
Figure 10. Pattern of prevailing wind direction according to 10° segments (2-times weighted) with contribution of adjacent segments. GEO3N project. 1990-2008.....	76
Figure 11. Flowchart of the selection of QRH for quality analysis in the Rhône-Alpes region, France. GEO3N project.....	190

List of tables

Table 1. Frequency (n) and percentage (%) of industrial sources and source-periods inventoried in the Rhône-Alpes region between 1990 and 2008, and mean and standard deviation (SD) of their annual dioxin emission intensity (in g-TEQ/year). GEO3N project.....	73
Table 2. Weighted kappa coefficients (wk) and 95% CI observed in Lyon and Le Bugey for 1996, 2002 and 2008 with a GIS-based metric including a distance decline of $1/d^2$ and a prevailing wind direction divided by 10° segments (2-times weighted) with contribution of adjacent segments.	76
Table 3. Presentation of the 7 articles selected through the literature review carried out to define the GIS-based metric.....	170
Table 4. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested, in Lyon, in 1996, 2002 and 2008.....	173
Table 5. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested, in Le Bugey, in 1996, 2002 and 2008. ..	174
Table 6. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested including stack height (h) and exhaust smoke velocity (v), in Lyon, in 1996, 2002 and 2008.....	175
Table 7. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested including stack height (h) and exhaust smoke velocity (v), in Le Bugey, in 1996, 2002 and 2008.....	176
Table 8. Frequency (n) and percentage (%) of missing data and outliers of the QRH analyzed in May 2015. GEO3N project.....	191
Table 9. Descriptive analysis of the lifetime residential addresses from the QRH (excluding missing data and outliers). GEO3N Project. (1/2).....	192
Table 10. Descriptive analysis of the lifetime residential addresses from the QRH (excluding missing data and outliers). GEO3N Project. (2/2).....	193
Table 11. Descriptive analysis of the lifetime occupational addresses from the QRH (excluding missing data and outliers). GEO3N Project.	194

List of personal publications

Binachon B, Dossus L, Danjou AMN, Clavel-Chapelon F, Fervers B. 2014. Life in urban areas and breast cancer risk in the French E3N cohort. *European Journal of Epidemiology* 29:743–751; doi:10.1007/s10654-014-9942-z.

Danjou A, Fervers B, Boutron-Ruault M-C, Philip T, Clavel-Chapelon F, Dossus L. 2015. Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort. *Breast Cancer Research* 17:39; doi:10.1186/s13058-015-0536-9.

Faure E, Danjou AMN, Carretero C, Clavel-Chapelon F, Dossus L, Fervers B. ‘Accuracy of two geocoding methods for Geographic Information System based exposure assessment in epidemiological studies’. *Submitted to Environmental Health on 04/10/2016*

Danjou AMN, Coudon T, Leffondré K, Lévêque E, Faure E, Praud D, Boutron-Ruault MC, Philip T, Dossus L, Clavel-Chapelon F, Fervers B. ‘Airborne Dioxin Exposure and Breast Cancer Risk in a case-control study nested within the French E3N Prospective Cohort: a GIS-based approach’. *In preparation*

List of oral communications

Danjou AMN. 2013. Exposition aux dioxines via l'alimentation et risque de cancer du sein dans la cohorte prospective E3N. SFSE (Société Française Santé et Environnement), Lyon, France.

Danjou AMN. 2014. Dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort. CENS (Centre Européen pour la Nutrition & la Santé), Lyon, France.

Danjou AMN. 2014. Exposition aux dioxines via l'alimentation et risque de cancer du sein dans la cohorte prospective E3N. 19^{ème} Journée Scientifique de l'EDISS, Ecole Doctorale Interdisciplinaire Science-Santé, Lyon, France.

List of poster-presentations

Danjou AMN. 2013. Dietary Exposure to Dioxins and Risk of Breast Cancer among Women from the E3N French Prospective Cohort. ISEE (International Society for Environmental Epidemiology), Basel, Switzerland.

Danjou AMN. 2014. Exposition aux dioxines via l'alimentation et risque de cancer du sein dans la cohorte prospective E3N. NACRe (réseau Nutrition Alimentation Cancer Recherche), Lyon, France.

Danjou AMN. 2014. Environmental exposure to dioxins and risk of breast cancer: the GEO3N Pilot study in the Rhône-Alpes region, France. Journée Scientifique du CLARA, Lyon, France.

Danjou AMN. 2014. Environmental exposure to dioxins and risk of breast cancer: the GEO3N Pilot study in the Rhône-Alpes region, France. ISEE (International Society for Environmental Epidemiology), Seattle, Washington, USA.

Danjou AMN. 2015. Environmental exposure to dioxins and risk of breast cancer: the GEO3N Pilot study in the Rhône-Alpes region, France. ECE (European Congress of Epidemiology), Maastricht, The Netherlands.

Danjou AMN. 2016. Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort. IARC (International Agency for Research on Cancer), Lyon, France.

Danjou AMN. 2016. Environmental dioxin exposure index and breast cancer risk in a case-control study nested within the French E3N prospective cohort: considering time of exposure in the risk estimate. IARC, Lyon, France.

CHAPTER I. INTRODUCTION

1. Scientific background

1.1. Breast cancer among women

1.1.1. Tumor pathology

The human breast comprises structural elements, such as connective tissue, fat, blood vessels and lymphatic tissue, and functional elements, including the mammary gland which is composed of lobules and ducts lined by epithelial cells. Carcinomas in the breast arise from benign lesions of epithelial origin in the mammary gland (Bodian 1993).

Mammary carcinogenesis is a multistage process that develops through a latency period of several years before the clinical onset of the disease. Tumor initiation involves accumulation of acquired (and sometimes inherited) gene mutations in a single cell that will be transmitted during cell division, leading to unregulated proliferation, loss of ability to balance cell division and cell death, and loss of differentiation of the progeny cells. The gene mutations include oncogenes that activate cell proliferation, tumor suppressor genes that inhibit cellular function and genomic stability genes involved in DNA repair. The gene mutations may be inherited or acquired as a result of viral infection, DNA damage caused by carcinogens (genotoxic or initiating agents) or random errors. Promotion refers to a preneoplastic stage during which initiated cells expand clonally. Subsequent tumor progression is driven by the accumulation of additional genetic mutations that causes abnormal structure, loss of contact inhibition, increased infiltration capacity and induction of neo-angiogenesis, and mediates the transition towards malignant tumor growth. Carcinomas *in situ* are confined to the epithelial layer of the breast, while invasive carcinomas infiltrate neighboring tissues, blood or lymph. Metastasis refers to the dissemination of malignant cells to another organ or tissue, leading to multiple tumor sites (Beckmann et al. 1997; Stewart et al. 2003).

Breast cancer is a heterogeneous disease that can be classified into several subtypes, in particular according to hormone receptors (Anderson et al. 2014). Approximately 70-80% of all breast tumors are composed of breast cells expressing estrogen receptors (ER) and progesterone receptors (PR) on their surface; these are called hormone-sensitive cancers. Estrogen and progesterone can then bind to the breast cells, influence the cell function and contribute to the proliferation of cancer cells (Keen and Davidson 2003). The detection of hormone receptors is an important indicator for the potential response to endocrine therapy.

1.1.2. Descriptive epidemiology

Worldwide

Cancer of the breast is the first cause of cancer among women, with approximately 1.67 million new cases diagnosed worldwide in 2012 representing 25% of all cancers (Ferlay et al. 2015). From 1980 to 2010, breast cancer incidence has increased worldwide by an estimated 3.1% annual rate. Incidence rates of breast cancer vary across world regions by a 1:4 ratio, with higher rates in more developed regions (Figure 1). It is estimated that breast cancer affects almost 1 in 8 women in the Western world (American Cancer Society 2015). While incidence rates are higher in developed countries, almost half of all breast cancer cases are diagnosed in developing countries (Forouzanfar et al. 2011). Nearly 2/3 of breast cancer cases are diagnosed in women aged 50 years and older, among which 58% live in developed countries. In contrast, more than 2/3 of new breast cancer cases in women aged 15-49 years are diagnosed in developing countries (Forouzanfar et al. 2011).

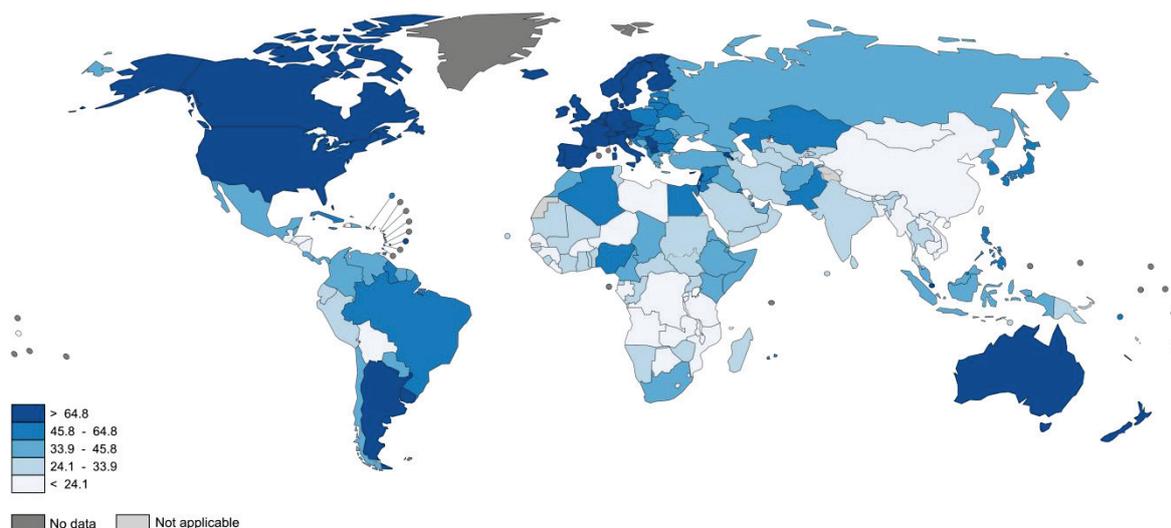


Figure 1. Estimated age-standardized female breast cancer incidence rates (per 100,000) worldwide in 2012. GLOBOCAN 2012, International Agency for Research on Cancer

Breast cancer is also the most frequent cause of death in women, with an estimated number of 522,000 deaths worldwide in 2012 (Ferlay et al. 2015). Mortality rates are higher in less developed countries where breast cancer is the most frequent cause of cancer death in women, while breast cancer represents the second cause of cancer death in more developed countries (Figure 2). The number of deaths from breast cancer worldwide has increased from 250,000 in 1980 to 425,000 in 2010, with an annual rate of 1.8% increase (Forouzanfar et al. 2011). Overall, mortality rates are lower than incident rates, due to the high survival after breast cancer in developed regions (Ferlay et al. 2015).

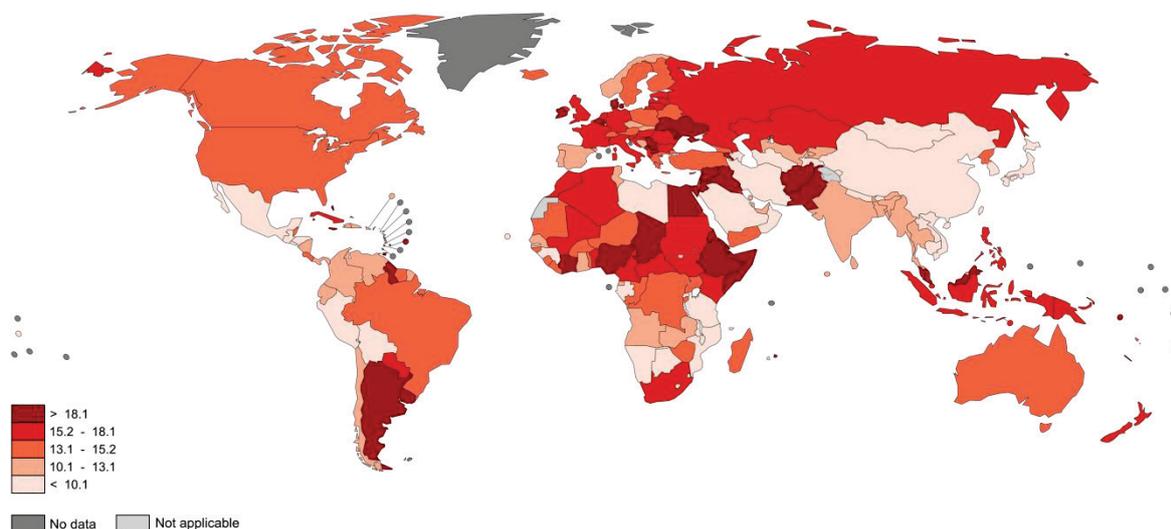


Figure 2. Estimated age-standardized female breast cancer mortality rates (per 100,000) worldwide in 2012. GLOBOCAN 2012, International Agency for Research on Cancer

The French context

In France, breast cancer is the most common malignant disease among women, with an estimated number of 54,062 new breast cancer cases and an age-standardized incidence rate of 94.7 per 100,000 predicted for 2015. For the past thirty years, the number of breast cancer cases has more than doubled, from 21,000 in 1980 to 49,000 in 2012 (Figure 3) (INCa 2015). The breast cancer incidence rate has increased until 2005 and, after a short decrease until 2008, has stabilized around 90 per 100,000

(Figure 3). The number of incident cases is the highest among women of 50 to 65 years of age (Figure 4).

A number of hypotheses has been put forward to explain the recent important increase in breast cancer incidence, such as ageing of the population and organized and individual breast cancer screening. Major changes in lifestyle in the past 50 years have also been pointed out, as demonstrated by studies on migrant populations for whom breast cancer incidence rates, and those in successive generations, align with the incidence of the host country (Kliwer and Smith 1995; McPherson et al. 2000). However, these factors do not explain entirely the observed changes in breast cancer incidence and the differences in the distribution of cancers observed between countries and world regions.

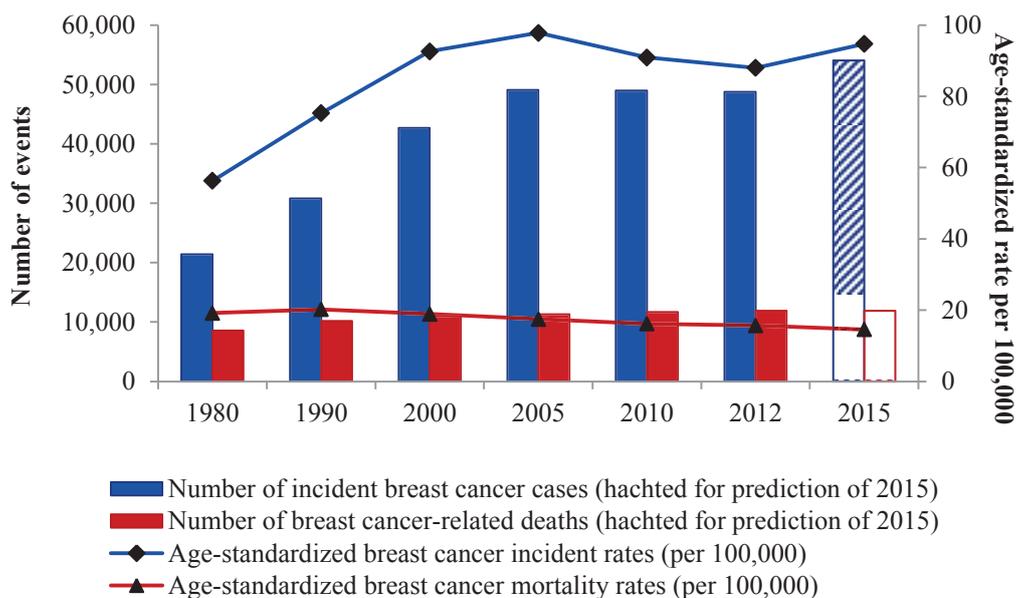


Figure 3. Evolution of the number of incident cases and breast cancer-related deaths between 1980 and 2015 (prediction). Evolution of age-standardized incidence and mortality rates (per 100,000) between 1980 and 2015. Institut National du Cancer, 2015

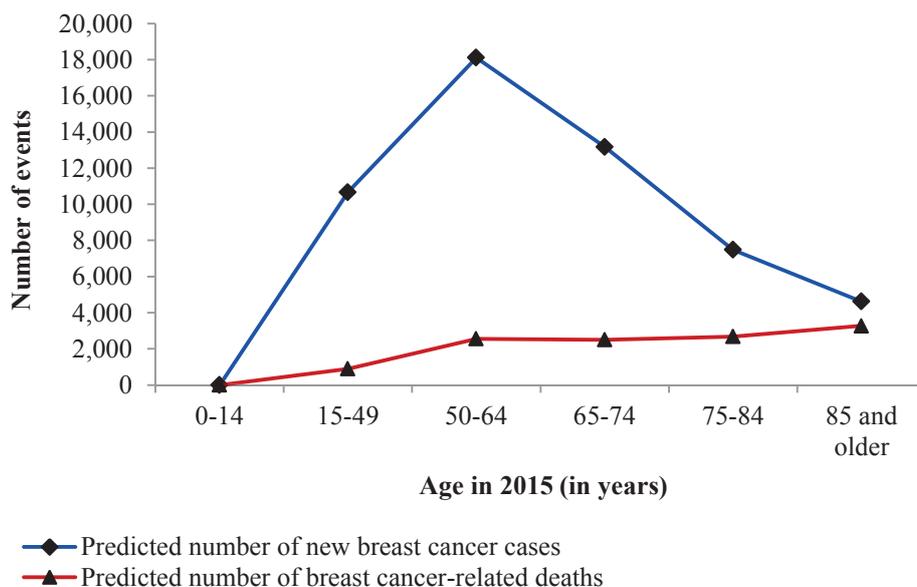


Figure 4. Number of new breast cancer cases and breast cancer-related deaths among women in France by age group – prediction for 2015. Institut National du Cancer 2015

Breast cancer is the leading cause of cancer-related death in women in France. In 2015, 11,913 breast cancer-related deaths have been predicted, representing an age-standardized mortality rate of 14.6 per 100,000 (INCa 2015). After an increase between 1980 and 1990, the mortality rate has been decreasing by 1% per year (Figure 3). Improvements in therapeutic practices and early detection due to the implementation of screening have certainly contributed to this decrease (Belot et al. 2008; Molinié et al. 2014). The number of breast cancer-related deaths increases strongly after 50 years of age, with 2,555 estimated breast cancer deaths in 2015 in women aged 50-64 years (Figure 4).

1.1.3. Etiology

The etiology of breast cancer is multifactorial and involves genetic, hormonal and reproductive, lifestyle and environmental factors. Numerous factors that influence the risk of breast cancer have been identified; however, the etiology of breast cancer remains in part unknown.

Age and sex

Sex is certainly the major risk factor for breast cancer. Breast cancer affects both men and women; however, the incidence is much higher in women and breast cancer in men represents less than 1% of all breast cancers in France (INCa 2015). Breast cancer incidence rates increase with age, exponentially before menopause and slower thereafter (Benz 2008).

Heritable factors

A history of breast cancer in first-degree relatives (mothers, sisters and daughters) is since long known to be a major risk factor for breast cancer. In a collaborative analysis of fifty-two epidemiological studies, the Collaborative Group on Hormonal Factors in Breast Cancer (CGHFBC) showed that the risk of breast cancer increased significantly with an increasing number of affected first-degree relatives. Risk ratios were 1.80 (99% confidence interval (CI)=1.69-1.91), 2.93 (2.36-3.64) and 3.90 (2.03-7.49) for respectively, one, two and three or more affected first-degree relatives (Collaborative Group on Hormonal Factors in Breast Cancer and others 2001).

Some women are at increased risk of breast cancer due to inherited genetic susceptibility. Mutations in breast cancer-susceptibility genes have been found at a high proportion in multiple case families, which are characterized by a family cluster of early-onset breast cancer cases and the inheritance of high-risk mutations (Thompson and Easton 2004). The two most important high-risk susceptibility genes that have been identified are BRCA1 (on chromosome 17q) and BRCA2 (on chromosome 13q), whose function is to maintain genome stability against DNA damage (Keen and Davidson 2003; Lee and Boyer 2001). In a pooled analysis of 22 studies, a 65% (95%CI=44%-78%) and a 45% (31%-56%) increases in lifetime breast cancer risk have been reported among BRCA1-mutation carriers and BRCA2-mutation carriers, respectively (Antoniou et al. 2003). Other mutations in susceptibility genes have been found, that moderately increase lifetime breast cancer risk, such as the CHEK2 gene (a kinase involved in the cell cycle at the G2 checkpoint, which has a critical role in DNA repair) and the ATM gene (a protein activated in response to DNA damage by ionizing radiation) (Thompson and Easton 2004). Approximately 5 to 10% of female breast cancers are attributed to inherited mutations in known susceptibility breast cancer genes (Apostolou and Fostira 2013; van der Groep et al. 2011).

An increasing number of breast cancer genome-wide association studies (GWAS) have been published, investigating the association between genetic variants, also called single nucleotide polymorphisms (SNP), and breast cancer. GWAS examine all or most of the genes in the genome of different individuals (at least 100,000 SNPs of hundreds or thousands of individuals) to identify the extent to which the genes vary from individual to individual (Peng et al. 2011). A review of genetic

polymorphisms and breast cancer risk reported results from 8 GWAS that included 25 SNPs significantly associated with increased breast cancer; the mean odds ratio (OR) for breast cancer was 1.19 and ranged from 1.04 to 1.43 (Peng et al. 2011). The most recent review reported that GWAS revealed 90 established breast cancer risk loci during the last decade that account for 14% of inherited breast cancer risk (Fachal and Dunning 2015).

Complex interactions between genetic and environmental factors can also influence breast cancer risk. Gene-environment interaction studies are particularly useful for identifying individuals who may be more susceptible to cancer, identifying novel genes through interactions and understanding biological pathways and mechanisms of disease etiology. Results from the Breast Cancer Association Consortium (BCAC, 34,793 invasive breast cancers and 41,099 controls, 23 SNPs) reported significant interactions between genetic variants of LSP1 (lymphocyte-specific protein 1, encodes for a binding protein) and parity and between genetic variants of CASP8 (caspase 8, apoptosis-related cysteine peptidase) and alcohol; suggesting that the risk of breast cancer associated with common genetic variants may vary according to environmental factors (Nickels et al. 2013). Conversely, the Breast and Prostate Cancer Cohort Consortium (BPC3) studied gene-environment interactions with 39 breast cancer risk SNPs and established breast cancer risk factors (16,285 breast cancer cases and 19,276 controls) and found no significant interaction, but a suggestive interaction between smoking status and a genetic variant of SLC4A7 (sodium bicarbonate cotransporter, mediates movements of sodium and bicarbonate ions across the plasma membrane) (Barrdahl et al. 2014). A recent review on gene-environment studies in cancer found that a majority of publications examined breast cancer (35%) and that specific interactions included energy balance, exogenous and endogenous hormones, chemical environment and lifestyle. Statistically significant breast cancer gene-environment interactions were reported in multiple publications for NAT2 (N-acetyltransferase 2, role in detoxification of aromatic monoamines) x lifestyle, XRCC1 (X-ray cross complementing group 1, DNA repair gene) x lifestyle, and MTHFR (methylene tetrahydrofolate reductase, role in building proteins) x energy balance. The review recommended exploring GWAS approaches in gene-environment interaction research (Simonds et al. 2016).

Benign breast conditions and breast density

Benign breast conditions are identified primarily among premenopausal women. A meta-analysis of 32 studies showed that benign breast disease was significantly associated with an increased risk of subsequent breast cancer, with summary risk estimates from 1.76 (95%CI=1.58-1.95) for proliferative disease without atypia to 3.93 (3.24-4.76) for atypical hyperplasia (Dyrstad et al. 2015). A diagnosis of carcinoma *in situ* has also been associated with a subsequent increased risk of invasive breast cancer (Bodian 1993).

It has been shown that percent mammographic density is associated with higher breast cancer risk. A meta-analysis of 42 studies found increased risks of breast cancer ranging from 1.79 (95%CI=1.48-2.16) to 4.64 (3.64-5.91) for women having 5-24% and $\geq 75\%$ breast density compared with women with little (<5%) or no breast density (McCormack and Silva 2006; Wang et al. 2014). Moreover, some breast cancer risk factors have also been associated with increased breast density (Vachon et al. 2000).

Hormonal and reproductive factors

The role of reproductive and menstrual factors in the etiology of breast cancer is well recognized, and there is strong evidence that hormones, in particular estrogens, both endogenous and exogenous, are involved in breast cancer development through both receptor dependent and independent mechanisms.

Menstrual factors relevant to the etiology of breast cancer include an earlier age at menarche (before 12 years old) and a late age at menopause (after 55 years old) that have been found associated with a greater risk of breast cancer (Bernstein 2002). Breast cancer risk has been shown to decrease by 4% (95%CI=2%-5%) to 9% (7%-11%) for each additional year in age at menarche (Clavel-Chapelon and Gerber 2002). Late age at menopause is associated with a greater breast cancer risk (relative risk (RR)=1.03, 95%CI=1.02-1.03, for each year older at menopause), due to a longer time period of exposure to ovulatory menstrual cycles (Collaborative Group on Hormonal Factors in Breast Cancer 1997).

Conversely, factors known to confer a reduced risk of breast cancer include parity and number of childbirths, earlier age at first full-term pregnancy, and breastfeeding. The CGHFBC found that parous women had an up to 30% reduced risk of breast cancer compared with nulliparous women (Clavel-Chapelon and Gerber 2002; Collaborative Group on Hormonal Factors in Breast Cancer and others 2002). Moreover, for parous women, the risk of breast cancer decreases with the number of children; but increases with the age at first full-term pregnancy. A recent meta-analysis of 100 studies showed that ever breastfeeding was associated with a 22% reduced breast cancer risk compared with never breastfeeding (OR=0.78, 95%CI=0.74-0.82). Breastfeeding for more than 12 months was found associated with a 26% lower risk of breast cancer compared with never breastfeeding (OR=0.74, 95%CI=0.69-0.79) (Chowdhury et al. 2015).

Factors related to an exogenous source of estrogens, usually taken in the form of medication, have been identified. For oral contraceptives, current and recent (1-4 years and 5-9 years after stopping), have been associated with higher breast cancer risk compared with never users, with RR=1.24 (95%CI=1.15-1.33), 1.16 (1.08-1.23) and 1.07 (1.02-1.13), respectively. There was no statistically significant increased breast cancer risk for 10 or more years after stopping use (RR=1.01, 95%CI=0.96-1.05) (Collaborative Group on Hormonal Factors in Breast Cancer and others 1996; Kumle et al. 2002). The effect of oral contraceptives on the risk of breast cancer decreases gradually after cessation of use, returning to that of never-users within 10 years after cessation (RR=1.01, 95%CI=0.96-1.05) (Collaborative Group on Hormonal Factors in Breast Cancer and others 1996). However, other results suggested no increase in breast cancer risk among oral contraceptive users compared with never users (Hannafoord et al. 2007; Marchbanks et al. 2002). Use of menopausal hormone therapies (MHT) has been convincingly shown as a risk factor of postmenopausal breast cancer. However, risks associated to the different types of MHT differed: results of the Women's Health Initiative trial (Rossouw et al. 2002) and of the Million Women Study cohort (Collaborators on the Million Women Study 2003) showed an increased risk with use of estrogen combined with synthetic progestins (RR=1.26, 95%CI=1.00-1.59 and RR=2.00, 95%CI=1.88-2.12, respectively); the E3N cohort (Etude Epidémiologique auprès de femmes de l'Education Nationale) study showed increased risks of breast cancer associated with the use of estrogen only (hazard ratio (HR)=1.29, 95%CI=1.02-1.65) and estrogen combined with a progestogen other than progesterone/dydrogesterone (HR=1.69, 95%CI=1.50-1.91) (Fournier et al. 2014). Women exposed *in utero* to diethylstilbestrol (DES), a synthetic estrogen that was administered to pregnant women to prevent miscarriage, have been shown at increased risk of breast cancer (HR=1.82, 95%CI=1.04-3.18) (Hoover et al. 2011), as well as mothers who took the medication (RR=1.4, 95%CI=1.1-1.9) (Greenberg et al. 1984). Moreover, the risk of breast cancer was higher for daughters whose mothers received the highest cumulative dose of DES during pregnancy, with incidence rate ratios of 1.63 (95%CI=0.87-3.08) for low-dose exposure *versus* 2.16 (1.18-3.96) for high-dose exposure (P-trend=0.01) (Palmer et al. 2006). These results suggest that an early hormonal exposure can affect breast cancer occurrence later in life.

Population-attributable fractions were used to estimate, in the E3N cohort cases proportions attributable to risk factors under hypothetical scenarios of lowest exposure. Regarding hormonal and

reproductive factors, results showed that age at menarche accounted for almost 20% of premenopausal and 10% of postmenopausal breast cancer, and that 15%, 7% and 5% of breast cancer could be avoided after menopause if women did not use MHT, had a menopause before 48 years of age, and more than one child – the first before age 30, respectively (Dartois et al. 2016).

Lifestyle factors

Several lifestyle-related factors are implicated in the etiology of breast cancer. They include anthropometric factors and personal behaviours, such as alcohol consumption, diet, smoking and physical activity.

The association between body mass index (BMI, as weight/height²) or body fatness (which includes overweight, obesity and abnormal or excessive accumulation of body fat (Lauby-Secretan et al. 2016)), and breast cancer risk differs by menopausal status. A meta-analysis of prospective studies highlighted opposite trends: a 5 kg/m² increase in BMI was found inversely associated with premenopausal breast cancer risk (RR=0.89, 95%CI=0.84-0.94) and positively associated with postmenopausal breast cancer risk (RR=1.09, 95%CI=1.04-1.14) (Latino-Martel et al. 2016; Renehan et al. 2008). Furthermore, studies have reported a positive association between taller height and breast cancer risk, among both premenopausal and postmenopausal women; however, the risk was of greater magnitude for postmenopausal women (pooled RR=1.02, 95%CI=0.96-1.10 and RR=1.07, 95%CI=1.03-1.12 per 5 cm height increment, respectively) (Friedenreich 2001; Van Den Brandt et al. 2000).

Alcohol consumption is also known to confer a higher breast cancer risk and studies found that consumption of large amounts of alcohol is associated with a higher risk of breast cancer (RR=1.32, 95%CI=1.19-1.45 for a 3 to 4-drink intake per day and RR=1.46, 95%CI=1.33-1.61 for more than 4 drinks per day, compared with women who reported no drinking), and suggested a dose-response relationship, with a 7.1% (5.5%-8.7%) increase in breast cancer risk for each additional daily drink of alcohol. Results have been consistent across types of alcohol consumed and similar for premenopausal and postmenopausal women (Collaborative Group on Hormone Factors and Breast Cancer 2002; Coronado et al. 2011; Seitz et al. 2012; American Institute for Cancer Research and World Cancer Research Fund 2010).

The International Agency for Research on Cancer (IARC) classified tobacco smoking as carcinogenic to humans (Group 1) with sufficient evidence in humans and recently reported a positive association between current tobacco smoking and breast cancer incidence (RR ranging from 1.1 to 1.3), which is both biologically plausible and consistent with causality (IARC 2012b). Reviews have also reported positive associations between breast cancer risk and active tobacco smoking (pooled RR=1.46, 95%CI=1.15-1.85) (Johnson 2005), as well as for duration (40 years *versus* 0: RR=1.50, 95%CI=1.19-1.89) and intensity (40 cigarettes per day *versus* 0: RR=1.20, 95%CI=1.00-1.44) of active tobacco smoking (Cui et al. 2006). A study conducted in the European Prospective Investigation into Cancer and Nutrition (EPIC) found an increase in breast cancer risk among current (HR=1.16, 95%CI=1.05-1.28) and former smokers (HR=1.14, 95%CI=1.04-1.25) and among women exposed to passive smoking (HR=1.10, 95%CI=1.01-1.20), compared with never smokers and women not exposed to passive smoking (Dossus et al. 2014). Moreover, analyses showed that breast cancer risk increased with increasing number of pack-years from menarche to first full-term pregnancy (HR=1.73, 95%CI=1.29-2.32 for every increase of 20 pack-years) (Dossus et al. 2014) and was higher in the premenopausal period (RR=1.68, 95%CI=1.33-2.12) (Johnson 2005).

The role of diet in the development of breast cancer has been investigated. An increase in postmenopausal breast cancer risk was associated with high consumption of fat (RR=1.32, 95%CI=1.11-1.58), and subtypes of fat: saturated fat (RR=1.13, 95%CI=1.05-1.22), monounsaturated

fat (HR=1.12, 95%CI=1.03-1.21) and polyunsaturated fat (HR=1.10, 95%CI=1.01-1.20) (Thiebaut et al. 2007). An association with high red meat intake was suggested (RR=1.17, 95%CI=1.06-1.29 (Boyd et al. 2003); RR=1.04, 95%CI=1.00-1.07 (Alexander et al. 2010)), in particular among premenopausal women (RR=1.24, 95%CI=1.08-1.42, (Taylor et al. 2009)) and for ER-positive and PR-positive tumours (RR from 1.42, 95%CI=1.06-1.90 for more than 5 servings of red meat per week to 1.97, 95%CI=1.35-2.88 for more than 1.5 servings per day, P-trend=0.001, (Cho E et al. 2006)). A meta-analysis of 15 prospective studies reported a reduction in breast cancer associated with high intake of fruits (RR=0.92, 95%CI=0.86-0.98) and fruits and vegetables combined (RR=0.89, 95%CI=0.80-0.99) (Aune et al. 2012b). Moreover, a recent pooled analysis of 20 prospective studies showed a decrease in ER-negative breast cancer risk associated with the consumption of vegetables (RR=0.82, 95%CI=0.74-0.90) (Jung et al. 2013). An inverse association between dietary fiber intake and breast cancer risk was reported in a recent meta-analysis of 16 prospective studies (RR=0.93, 95%CI=0.89-0.98), and the association appeared to be most pronounced with high intakes (RR=0.91, 95%CI=0.86-0.97 for ≥ 25 g/day) (Aune et al. 2012a; Latino-Martel et al. 2016). Findings of a recent meta-analysis were in favour of a decrease in breast cancer risk with total dairy products intake (RR=0.85, 95%CI=0.76-0.95) (Dong et al. 2011). Studies also focused on dietary patterns, showing a protective effect of a “Prudent/Healthy” diet rich in raw vegetables and olive oil (OR=0.89, 95%CI=0.82-0.99) (Brennan et al. 2010) and increased breast cancer risks with the “Western” pattern characterised by high consumption of alcohol, meat products, French fries, rice/pasta, butter/cream, etc. (HR=1.20, 95%CI=1.03-1.38, (Cottet et al. 2009)) and the “Drinker” dietary pattern (OR=1.21, 95%CI=1.04-1.41, (Brennan et al. 2010)).

Regular physical activity has been associated with decreases in premenopausal breast cancer risk (RR=0.77, 95%CI=0.72-0.84), postmenopausal breast cancer risk (RR=0.88, 95%CI=0.84-0.92) and ER-negative/PR-negative breast cancer risk (RR=0.80, 95%CI=0.73-0.87). Moreover, studies have found a dose-response relationship showing a 2% to 5% decrease in breast cancer risk with increasing levels of physical activity (American Institute for Cancer Research and World Cancer Research Fund 2010; Wu et al. 2013).

Occupational factors

Few occupational factors have been identified as breast cancer risk factors. Studies have investigated the risk of breast cancer among night shift workers. The analysis of the U.S. Nurses' Health Study found that nurses had an elevated risk of breast cancer after long periods of rotating night shift-work (RR=1.79, 95%CI=1.06-3.01 for more than 20 years), compared with nurses who did not report any rotating night shift-work; the association did not remain statistically significant for fewer years of rotating night shift (Schernhammer et al. 2006). In 2007, IARC classified “shift-work that involves circadian rhythm disruption” as probably carcinogenic to humans (Group 2A), on the basis of sufficient evidence in experimental animal studies and limited evidence in humans (Straif et al. 2007). A recent meta-analysis concluded to a positive association between circadian disruption and increased breast cancer risk in women (RR=1.14, 95%CI=1.08-1.21) (He et al. 2014).

Environmental factors

In this section, the term “environmental” refers to contaminants such as chemicals, pollutants and radioactive substances whose exposure may occur from the indoor and outdoor air, food, water and soils. Few environmental risk factors for breast cancer have been established; others are suspected to increase breast cancer risk.

Ionizing radiation is an established cause of breast cancer and has been classified as carcinogenic to humans (Group 1) by IARC (El Ghissassi et al. 2009; IARC 2012a). X-rays and gamma-rays are

ionizing electromagnetic radiations that penetrate living tissues. Epidemiological studies showed that women are at higher risk of breast cancer when exposed to radiations related to diagnosis (OR=10.6, 95%CI=1.93-58.0, for a history of 3 to 5 chest X-rays (John et al. 2013)) or treatment of health conditions (RR=1.85, 95%CI=1.14-3.00 for medical radiation exposure (Hung and Hwang 2013)), and from radioactive substances in occupational settings (OR=2.90, 95%CI=1.19-7.04 (Wang et al. 2015)). There is no evidence of an association between exposure to electromagnetic fields and risk of breast cancer (Feychting and Forssén 2006).

Numerous environmental pollutants have been shown, in animal studies, to damage DNA, alter mammary gland development and promote tumor development and growth; 216 chemicals have been identified (Rudel et al. 2007, 2014). Brody et al. reviewed epidemiological studies that investigated breast cancer risk in relation to exposure to mammary gland carcinogens, in particular persistent organochlorine compounds (Brody et al. 2007).

Findings on the association between risk of breast cancer and polychlorinated biphenyls (PCB) in the general population were inconsistent, regardless of the exposure measures and for individual congeners. While most case-control studies reported no associations between breast cancer risk and total/individual PCBs measured in serum or adipose tissue, a few other found statistically significant increased risks of breast cancer associated with high PCB serum or adipose concentrations, also for total PCB and individual congeners (Brody et al. 2007). PCBs were classified as carcinogenic to humans (Group 1) by IARC in 2013, and some statistically significant positive associations have been observed for breast cancer (Aronson et al. 2000; Demers et al. 2002; Millikan et al. 2000; Moysich et al. 1998, 1999). However, although the associations may be biologically plausible, they were considered as providing limited evidence due to inconsistent results. Inconsistency may be explained by imprecise findings, small study population, residual confounding and the selection of controls with benign breast disease; no association was reported between occupational exposure to PCBs and breast cancer risk (Lauby-Secretan et al. 2013; IARC 2016). Recently, a meta-analysis of 16 studies investigating correlation between PCB congeners and breast cancer found a statistically significant increase in breast cancer risk among individuals with higher plasma and fat levels of PCB 99 (pooled OR=1.36, 95%CI=1.02-1.80), PCB 183 (OR=1.56, 95%CI=1.25-1.95) and PCB 187 (OR=1.18, 95%CI=1.01-1.39) (Leng et al. 2016).

Pesticides have received some attention with regard to breast cancer risk in women. A recent meta-analysis of 35 studies (16 hospital-based case-controls studies, 11 population-based case-control studies and 10 nested case-control studies) found no evidence of an increase in breast cancer risk associated with dichlorodiphenyltrichloroethane (DDT, measured in serum, plasma and adipose tissue) in the recent published literature: the summary OR was 1.03 (95%CI=0.95-1.12). However, study limitations were reported such as the delay between DDT exposure and diagnosis, age of exposure and combined exposure with other carcinogens (Park et al. 2014). Results for other organochlorine pesticide exposures, such as dieldrin and hexachlorobenzene, and breast cancer risk are not conclusive and further research is needed (Brody et al. 2007; Xu et al. 2009).

Major sources of polycyclic aromatic hydrocarbon (PAH) exposure for the general population are tobacco smoke, air pollution, vehicle exhaust (including diesel) and smoked and grilled foods. Recently, household air pollutants from indoor combustion, which releases PAHs, have been positively associated with breast cancer. White et al. reported an increased risk of breast cancer for ever burning synthetic logs indoor (OR=1.42, 95%CI=1.11-1.84) and the association was even stronger among women reporting more than 7 years of use (OR=1.73, 95%CI=1.11-2.70) and among women over 20 years of age (OR=1.65, 95%CI=1.02-2.67) (White et al. 2014). Vehicular traffic PAHs have been positively associated with breast cancer incidence in the Long Island Breast Cancer Study Project (LIBCSP): a first report among 575 breast cancer cases and 424 controls showed that PAH-

DNA adducts were positively associated with breast cancer (age-adjusted OR=1.51, 95%CI=1.04-2.20, for the highest vs the lowest quintiles), however the multivariate-adjusted OR was borderline significant (OR=1.49, 95%CI=1.00-2.21) (Gammon et al. 2002a). A second report of the LIBCSP examined the association between residential exposure to vehicular traffic and breast cancer incidence, among 1508 breast cancer cases and 1556 controls. Overall, a non-significant increase of breast cancer risk was associated with long-term vehicular traffic estimates (OR=1.44, 95%CI=0.78-2.68) (Mordukhovich et al. 2015). In the same study population as above, exposure to multiple sources of PAHs was associated with increased breast cancer risk: OR=1.20 (95%CI=1.03-1.40) for residential environmental tobacco smoke; OR=1.29 (1.06-1.57) for synthetic firelog burning and OR=1.45 (1.02-2.04) for total indoor sources including active smoking, environmental tobacco smoke from spouse, grilled/smoked meat intake, and stove/fireplace use (White et al. 2016).

Ethylene oxide, which is a sterilizer, was classified as carcinogenic to humans (Group 1) in 2009 with limited evidence for humans and strong mechanistic evidence due to its genotoxicity. However, findings for breast cancer were not consistent across studies (Baan et al. 2009): three cohort studies of workers exposed to ethylene oxide found overall no excess in breast cancer risk (Hagmar et al. 1995; Steenland et al. 2003) and a borderline significant 60% excess risk (Norman et al. 1995). A high level of cumulative exposure to ethylene oxide was found associated with increased risk of breast cancer (OR=1.87, 95%CI=1.12-3.10 for a 15-year lag) adjusted for parity and familial history of breast cancer (Steenland et al. 2003).

An increase in breast cancer risk was found associated with environmental exposure to phthalates in a case-control study (233 breast cancer cases and 221 controls): phthalates metabolites were detected in 82% of women; urinary monoethyl phthalate (the main diethyl phthalate metabolite) concentration was positively associated with breast cancer risk after adjusting for risk factors (OR=2.10, 95%CI=1.33-3.63, p-trend <0.01, for the highest vs the lowest tertile) and the increase was stronger among premenopausal women (OR=4.13, 95%CI=1.60-10.10, p-trend <0.01, for the highest vs the lowest tertile) (López-Carrillo et al. 2009). A role of phthalates at low concentration (Chen and Chien 2014) and in the promotion of tumor growth of ER-negative breast cancer cells has also been demonstrated (Giulivo et al. 2016; Hsieh et al. 2012).

1.2. Environmental exposure to air pollutants with endocrine disruptive properties

1.2.1. Air pollutants with endocrine disruptive properties

As defined by the U.S. Environmental Protection Agency (EPA), endocrine-disrupting compounds (EDC) are “exogenous agents that interfere with synthesis, secretion, transport, metabolism, binding action, or elimination of natural blood-borne hormones that are present in the body and are responsible for homeostasis, reproduction, and developmental process” (Diamanti-Kandarakis et al. 2009; Kavlock et al. 1996). EDCs represent a broad class of heterogeneous molecules, including organochlorine pesticides, industrial chemicals and their by-products, plastics and plasticizers, heavy metals, fuels, and many other chemicals. They are released from industrial and commercial uses into the environment, where they persist due to their long half-life. Exposure to EDCs occurs through breathing contaminated air, contacting contaminated soil, drinking contaminated water and ingesting contaminated food. They are found in many products including plastic bottles, metal food cans, detergents, flame retardants, food additives, toys, cosmetics and pesticides (Schug et al. 2011). There is evidence that EDCs have effects on breast development by altering mammary gland morphogenesis (Diamanti-Kandarakis et al. 2009). The mechanisms of EDCs involve in particular the estrogenic pathway, through which they can (i) mimic the effect of endogenous hormones such as estrogens and androgens, (ii) antagonize the effect of endogenous hormones, and (iii) disrupt synthesis and

metabolism of endogenous hormones and (iv), modify hormone receptor levels (Sonnenschein and Soto 1998). Epidemiological studies and laboratory findings suggest that environmental pollutants with endocrine-disrupting properties may contribute to the development of hormone-dependent cancer such as breast cancer (Brody et al. 2007; Teitelbaum et al. 2015).

In 2013, IARC classified outdoor air pollution and particulate matter from outdoor air pollution as carcinogenic to humans (Group 1), with sufficient evidence in humans and experimental animals and strong mechanistic evidence for lung cancer (Loomis et al. 2013). Outdoor and indoor air pollutions are also suspected to be a risk factor for breast cancer. Air pollution is a mixture of multiple pollutants and include carcinogens with endocrine-disrupting properties classified individually by IARC as carcinogenic (Group 1, including PCBs, benzo[a]pyrene (B[a]P), cadmium and 2,3,7,8-tetrachlorodibenzo-*para*-dioxin (TCDD)), and probably carcinogenic (Group 2A, including furans) to humans (Baan et al. 2009).

1.2.2. Dioxins

Polychlorodibenzo-*para*-dioxins (PCDD) and polychlorodibenzofurans (PCDF), commonly named “dioxins”, are chlorinated PAHs. There are 210 identified dioxin congeners (75 PCDDs and 135 PCDFs) distinguished by the number and position of the chlorines atoms (between 4 and 8) that they possess. Among these congeners, 17 (7 PCDDs and 10 PCDFs) have been identified as toxic by the World Health Organization (WHO) and TCDD is the most toxic congener.

In 1997, IARC classified TCDD as carcinogenic to humans – Group 1, classification that was confirmed in 2009, with sufficient evidence in humans for all cancers combined and limited evidence for lung cancer, soft-tissue sarcoma and non-Hodgkin lymphoma (Baan et al. 2009; IARC 1997). TCDD acts as an endocrine disruptor *via* its binding and activation of the aryl hydrocarbon receptor (AhR), a ligand-activated transcription factor expressed in most mammalian tissues and conserved across species (Grassman et al. 1998; Jenkins et al. 2012; Safe 1995). TCDD-binding to AhR leads to the nuclear translocation of the dimer, heterodimerization with the AhR nuclear translocator (Arnt), which together activate transcription of genes with xenobiotic response element (XRE) or dioxin response element (DRE) at their regulatory site. The target genes include genes of the detoxication system (CYP1A1, CYP1A2 CYP1B1 and CYP2S1), and genes involved in the cellular response (proliferation and apoptosis) (Guo et al. 2009) (Figure 5: a.).

The primary mechanism of action of TCDD for carcinogenesis is the promotion of tumor development by activation of cellular replication, alteration in cellular senescence and apoptosis. The second mechanism is the increase in oxidative stress. TCDD is not directly genotoxic but causes an increase in reactive oxygen species (ROS) leading to DNA damage and mutations (Mandal 2005). Furthermore, ER binds to estrogen response elements (ERE) and activates transcription in an estrogen-dependent way (Figure 5: b.). Studies suggested that dioxins disrupt estrogenic regulations by binding to the AhR and may have differential effects depending on hormone levels (Brunnberg et al. 2011; M Hall 2013; Safe et al. 2013; Soto and Sonnenschein 2010). Dioxins appear to have anti-estrogenic responses in the presence of estrogens *via* inhibition of estrogen-induced genes and proteins (Krishnan et al. 1995), and estrogenic effects in the absence of estrogens by which dioxins activate transcription of hormone-dependent genes (Figure 5: d.) (Ohtake et al. 2003).

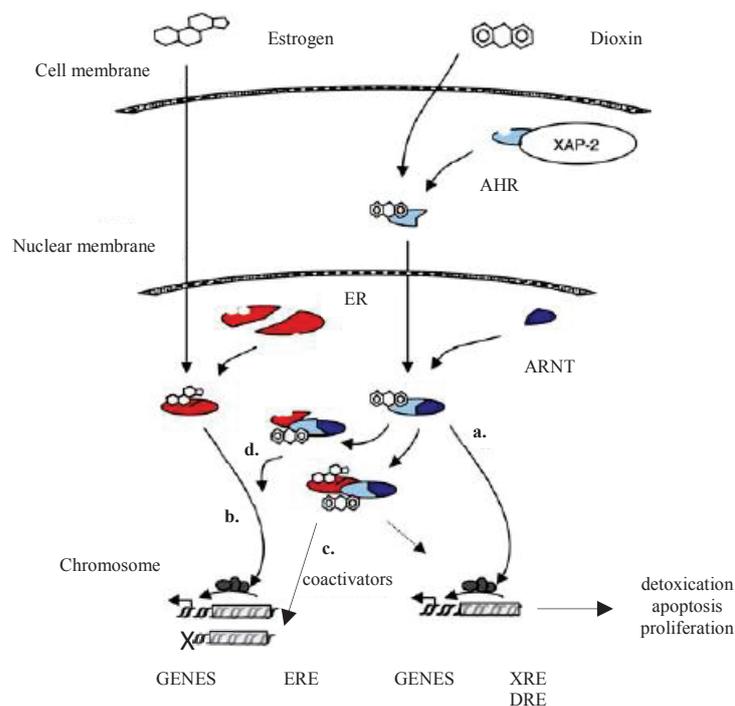


Figure 5. Dioxin-estrogen pathway crosstalk. Adapted from Guo 2009.

Chemically stable, TCDD is a persistent organic pollutant (POP) with a long half-life in the environment and in humans (estimated at 7.2 years (Milbrath et al. 2009)), and bio-accumulate through animal and human food chains. TCDD has low solubility in water, but is highly lipophilic, it accumulates in adipose tissue. TCDD has been detected in humans in adipose tissue, blood and breast milk (Arisawa et al. 2003; Sonawane 1995). Human exposure to dioxins is calculated in terms of toxic equivalence quotients (TEQ). Each of the 17 toxic dioxin congener is assigned a toxic equivalence factor (TEF), depending on its affinity with the Ah receptor, which is weighted relatively to TCDD. The TEQ of a mixture of congeners is calculated by multiplying individual levels of congeners by its TEF and summing the individual values.

Dioxin emissions

Dioxins are produced unintentionally during combustion processes involving a chlorine source. Sources of environmental release of dioxins have been grouped into four major categories (IARC 2012c):

- Emissions from incineration sources: municipal waste, hospital waste, hazardous waste and sewage sludge incineration;
- Emissions from combustion sources: cement kilns, wood-burning including domestic heating, diesel vehicles, coal-fired utilities and crematoria;
- Emissions from industrial sources: pulp and paper mills, chemical manufacturing and metal industry;
- Emissions from reservoir sources: biochemical processes, photolytic processes, forest fires, volcanic eruptions and accidental releases. These natural emissions are occasionally important but represent a very small proportion of the total emissions.

Dioxin emissions have evolved over the years because of changes in major emission sources and regulations. In the 1960s, the manufacture and use of chlorinated compounds (chlorophenols and chlorinated phenoxy herbicides) were the major sources of dioxin-release into the environment. These

emissions have been greatly reduced due to changes in legislation and the restricted use of organochlorine pesticides in agriculture. Then, incineration and combustion processes became the main sources of dioxin emissions. France possesses the largest park of Municipal Solid Waste incinerator (MSWI) in the European Union. Since 1995, the park has undergone a major transformation, mainly due to changes in legislation, in particular the ministerial order of September 20th 2002 (modified on August 3rd 2010 and August 22th 2016) that transposed into national law the European directive of December 4th 2000 (JORF n°280). Adapting the technologies of incineration plants to the changes in regulations led to a decrease in the number of facilities and an improved control of emissions (Anzivino-Viricel et al. 2012). In France, according to the Interprofessional technical centre for studies on air pollution (CITEPA), emissions of dioxins have decreased by 90% between 1990 and 2008, from 1776 g-TEQ estimated in 1990 to 143 g-TEQ estimated in 2008 (CITEPA 2015).

Human exposures

Dioxins are ubiquitous contaminants in the environment, including ambient air, surface water, groundwater, soil and sediment. Dioxin exposure in the general population is considered to be low; a review estimated the mean TCDD blood concentration at 1.8 pg/g lipid (standard deviation SD=1.7) in the general population not directly exposed to dioxin emitting sources (109 studies published in 1989-2010 (Consonni et al. 2012)). However, as dioxins are persistent pollutants and endocrine disruptors, the general population undergoes long-term exposure to a substantial amount, and dioxin blood levels continuously increase with age (Consonni et al. 2012).

IARC estimated that approximately 90% of dioxin exposure in the general population occurs through ingestion of contaminated food, as a result of the accumulation of these compounds in the food chain and in high-fat foods such as dairy products, eggs, fats and fish (IARC 2012c). Exposure through inhalation represents a small fraction of the human dioxin exposure. Some populations may be contaminated through the consumption of animals living – and foods produced – in the vicinity of industrial plants, and may be thus exposed at higher concentrations. In addition, some populations may have higher exposure due to proximity to specific sources of exposure.

In contrast to the general population, specific groups may have been exposed to dioxins at higher levels as a result of occupational exposure, in particular during production and use of chlorophenols and chloro phenoxy herbicides. Finally, even higher exposures have occurred because of industrial accidents, the most known being the explosion of the chemical plant Icmesa in Seveso (Italy, July 1976) that led to the release of a large amount of TCDD of up to 30 kg over 18.1 km² in the most contaminated area (di Domenico et al. 1990).

Due to their potential carcinogenicity and estrogen mimetic properties demonstrated in experimental studies, dioxins are suspected to be involved in breast cancer etiology (Brody et al. 2007; Kogevinas 2001).

1.3. Dioxin exposure and breast cancer risk

Several epidemiological studies have investigated the relationship between dioxin exposure and the risk of breast cancer. Evidence regarding dioxin exposure and breast cancer comes from populations accidentally and occupationally exposed or living in the vicinity of industrial facilities.

Accidental exposure

In 1976 in Seveso, Italy, an industrial accident exposed the local population after the release of a large amount of TCDD. The exposed population was classified into three categories according to soil

measurements: zone A (highest exposure), zone B (median exposure) and zone R (lowest exposure). Blood samples were collected between 1976 and 1981. A long-term cohort study was recruited that involved the population aged 0 to 75 years living in one of the three contaminated zones. First results, after a 10-year follow-up (1977-1986), found no association with breast cancer occurrence among women under and over 45 years of age (RR=0.9, 95%CI=0.3-2.9 and RR=0.7, 95%CI=0.3-1.5, respectively), compared with women living in a surrounding non-contaminated area (*i.e.* the reference area) (Bertazzi et al. 1993). After a longer follow-up period (1977-1996), a nonsignificant increase in breast cancer risk was observed among women from zone A compared with women from the reference area (RR=1.43, 95%CI=0.71-2.87); results remained similar in subgroups defined by menopausal status. However, in analyses restricted to breast cancer cases diagnosed after 15 years of follow-up, a significant increased breast cancer risk was found (5 cases, RR=2.57, 95%CI=1.07-6.20) (Pesatori et al. 2009). A specific retrospective cohort was conducted twenty years after the accident with 981 women aged 0 to 40 years at the time of the accident, living in zones A and B having provided blood samples after the accident. Results showed a dose-response relationship, with a 2-fold increase in breast cancer risk for a 10-fold increase in TCDD serum concentrations (15 cases, HR=2.1, 95%CI=1.0-4.6) (Warner et al. 2002). However, at that time, women were still young for breast cancer diagnosis and the most recent report from this cohort did not confirm the previous results (33 cases, HR=1.44, 95%CI=0.89-2.33) (Warner et al. 2011).

Occupational exposure

Epidemiological studies have been conducted among herbicide and pesticide female workers exposed to dioxins and suggested elevated breast cancer mortality. In a study combining data from 36 occupational cohorts from 12 countries, an overall non-statistically significant increased mortality was observed for the 1939-1992 period. A statistically significant increased mortality from breast cancer was observed when confined to one German cohort (9 deaths, standardized mortality ratio (SMR)=2.84, 95%CI=1.30-5.39) (Kogevinas et al. 1997). Recently, Manuwald et al. examined the long-term effects of dioxin exposure among former employees of a chemical plant in Hamburg, Germany. Twenty-three years after the closure of the plant, a significant increase of breast cancer mortality was found among women (19 deaths, SMR=1.86, 95%CI=1.12-2.91) (Manuwald et al. 2012).

Local exposure

An elevated breast cancer mortality was also reported among women in Chapaevsk, Russia, living near a chemical plant, in a contaminated area where dioxins had been detected in air, soil, tap water and cow milk (58 breast cancer-related deaths, SMR=2.1, 95%CI=1.6-2.7, compared with women from a surrounding non-contaminated area) (Revich et al. 2001). In Michigan, USA, a case-control study was conducted to assess dioxin soil concentration and explore the potential association with breast cancer incidence observed in the area. The study area involved three counties, in which there was a chemical plant. A significant spatial correlation was found between the increase in breast cancer incidence and the soil contamination by dioxins in the three areas compared with the reference area ($p < 0.05$). Moreover, a cluster of breast cancer cases was identified located in close proximity to the contaminated area (Dai and Oyana 2008).

In France, a national ecological study was conducted to investigate the relationship between cancer incidence and airborne emissions from MSWI in the general population. A statistically significant positive association was observed between living under the plume of a MWSI and the occurrence of breast cancer ($p = 0.03$). Moreover, a 10% statistically significant increased risk of breast cancer was found when comparing the 90th versus the 2.5th exposure percentiles (18,824 cases, RR=1.09,

95%CI=1.01-1.18) (Fabre et al. 2008). However, there was no adjustment for breast cancer risk factors in the analyses and occupational and dietary exposures were not considered. Another French population-based case-control study examined the association between airborne dioxin emissions from MSWI and risk of breast cancer, among women living under the plume of the MWSI. Results reported a decreased risk of breast cancer among women over 60 years and highly exposed to dioxins compared to women with the lowest dioxin exposure (434 cases, OR=0.31, 95%CI=0.08-0.89); analyses were not adjusted for breast cancer risk factors and other dioxin exposure routes (Viel et al. 2008).

Direct measurements among surgical patients

Two hospital-based case-control studies conducted among surgical patients investigated breast cancer risk in relation to several dioxin-congener concentrations measured in adipose tissue of the breast. Overall no association was reported with invasive breast cancer *versus* benign breast disease (Hardell et al. 1996; Reynolds et al. 2005). However, an increased risk of breast cancer was found associated with octachlorodibenzo-*p*-dioxin (OCDD) after adjustment for breast cancer risk factors (22 cases, OR=1.1, 95% CI=1.0-1.3) (Hardell et al. 1996). Results of the two studies were based on small populations, and included controls with benign breast conditions that may be at higher breast cancer risk than a non-hospital population. Moreover, the exposure measurements occurred near time of diagnosis and may not represent the relevant etiologic period.

1.4. Limitations and issues of existing studies on dioxin exposure and breast cancer

Results on the association between environmental exposure to dioxins and risk of breast cancer remain inconsistent (Brody et al. 2007). There is still a controversy whether or not dioxins increase breast cancer risk in the general population. Previous studies presented several limitations that need to be highlighted.

A major limitation is the inadequate control for confounding by individual breast cancer risk factors in statistical analyses. Also, most studies lacked information to stratify on the breast cancer hormone receptor status (ER and PR) although there is evidence that the association with air pollutants may differ according to hormone receptor status (Garcia et al. 2015; Liu et al. 2015; Raaschou-Nielsen et al. 2005). Except for the French ecological study, results were reported for small population sizes and small number of breast cancer cases. While diet is the primary route for dioxin exposure, to our knowledge, no study estimated the risk of breast cancer associated with dietary dioxin exposure (Brody et al. 2007).

The lack of past residential history of the study subjects and historical exposure estimates, the consideration of exposure from MSWI only and not from other facilities (metal and cement industries, traffic) and the assessment of airborne exposure based only on proximity to the studied facilities and neglecting the influence of the local meteorology and topography limited the interpretation of results and may have resulted in exposure misclassification, likely to contribute to risk estimates biased toward unity (Basagaña et al. 2013; Copeland et al. 1977). These issues suggest that the methodology for the assessment of retrospective environmental dioxin exposure needs to be improved (Brody et al. 2007; Cordioli et al. 2013).

Low-dose exposure

The current evidence of the carcinogenicity of TCDD is based on studies involving occupational and accidental populations (IARC 2012c). Exposure to dioxins was found associated with an increased risk of breast cancer in these populations, exposed to high levels of dioxins. These results cannot be extrapolated to the general population, which is routinely exposed to low-levels of dioxins, especially

because the effects of high doses of EDCs on tumor development may not be representative of the effects happening at low doses (Popp 2006). The American Cancer Society reported that there should be concerns about low-level exposure to carcinogenic pollutants, in particular because of the involuntary and long-term chronic nature of exposures and the potentially large population that can be exposed (Fontham et al. 2009).

Timing of exposure

The complex effects of timing of exposure and its assessment may further contribute to the inconclusive results observed in epidemiological studies. A too short latency period does not allow to observe of the effects of exposure; increasing time of follow-up is therefore recommended. Also measuring the exposure at the time of diagnosis or interview may not be representative of the etiological period for breast cancer.

Age at menarche, parity and age at first full term pregnancy are factors known to influence the risk of breast cancer over a life-time. The breast is not fully developed at birth and differentiates through puberty, pregnancy and lactation, leading to portions of the lifetime during which the mammary gland would be at particular vulnerability to exposure (Jenkins et al. 2012; Teitelbaum et al. 2015). In previous studies, it was not possible to differentiate these specific periods of exposure. Moreover, early life exposure to PCDD/F, including prenatal and lactational exposures, has been found associated with a delay in the initiation of breast development among girls (Den Hond et al. 2002), and in particular among those living near waste incinerator and metal industry (Leijs et al. 2008). Animal studies also demonstrated that prenatal dioxin exposure doubled the incidence of mammary tumors later in life and decreased latency (Fenton and Birnbaum 2015). Therefore, women may experience multiple time-variable windows of breast susceptibility in which dioxins could affect hormonal pathways and induce the development of breast cancer later in life (Lamartiniere 2002; Teitelbaum et al. 2015). However, there is a lack of knowledge about the breast cancer latency period, which is the time interval between dioxin exposure and breast cancer development. Given the tumor-promoter properties of dioxins, studying the association between dioxin exposure and risk of breast cancer according to specific windows of exposure should not be neglected, nor should the impact of late exposures.

In epidemiological studies, analysis of the exposure-response association can be challenging due to the long-term nature of the exposure that implies variations of exposure intensities over time and consideration of specific timings of exposure (Vacek 1997). Environmental exposure to dioxins is characterized by its bioaccumulation and persistence in the environment. Dioxin exposure has decreased in France for the last twenty years and, besides lifestyle factors (dietary habits and heating), individual dioxin exposure level is also related to the residential history of the population. Statistical methods have been developed that improve modeling of long-term and time-varying exposures, by using weighted functions according to the timing of exposure (Hauptmann et al. 2000; Sylvestre and Abrahamowicz 2009); their use in epidemiological studies is now strongly encouraged (Vacek 1997).

2. Objectives of the thesis

2.1. Objectives

The aims of the thesis were to investigate the associations between dietary and airborne dioxin exposures and breast cancer risk in women, improving the methodology of assessment of the low-level dioxin exposures and using adequate statistical methods to consider temporal dimensions of exposure in the risk estimate. Furthermore, particular attention was paid to hormonal receptor status (ER and

PR) and specific windows of susceptibility (menarche, pregnancy and menopause). The specific research objectives were:

- To assess the dietary dioxin exposure, using data on food consumption and food contamination by dioxins, and estimate the risk of breast cancer in the E3N prospective cohort (CHAPTER II).
- To develop a Geographic Information System (GIS)-based exposure metric to assess individual airborne dioxin exposure among women from the E3N cohort living in the Rhône-Alpes region (CHAPTER III).
- To estimate the risk of breast cancer associated with airborne dioxin exposure and the relative weight of the exposure in the association in the E3N cohort (CHAPTER IV).

From these research objectives, three scientific articles have been produced and published or submitted to international journals. The first one, presented in CHAPTER II, investigates the association between dietary dioxin exposure and breast cancer risk, using a method proposed by WHO for the assessment of dietary exposure and the estimation of the breast cancer risk in the E3N cohort (article published in *Breast Cancer Research*). The second article, presented in CHAPTER III, focuses on the geocoding method of the residential history of the E3N women of the Rhône-Alpes region, whose positional accuracy is a major factor to consider in epidemiological studies to limit misclassification (article submitted to *Environmental Health*). The third article, presented in CHAPTER IV, describes the GIS-based methodology developed for the assessment of the individual airborne dioxin exposure and the statistical methods applied to examine the relationship between airborne dioxin exposure and breast cancer risk. Preliminary results are presented among women from the case-control study nested within the E3N cohort and living in the Rhône-Alpes region (article in preparation).

2.2. Thesis context

The Centre Léon Bérard (Lyon, France) is a regional French Comprehensive Cancer Center that fulfill three main missions: care, research and teaching (Centre Léon Bérard 2016). In 2009, an interdisciplinary program on “Cancer, Environment and Nutrition” was implemented within the Centre Léon Bérard, leading to the creation of the Cancer and Environment Research Department. The research program of the Department focuses on:

- The investigation of environmental and occupational risk factors associated with breast cancer, testicular cancer and ovarian cancer through several research projects;
- The evaluation and nutritional management of breast cancer patients overweight or at risk of being and the introduction of a program of adapted physical activity for all cancer patients.

The Department is involved in the improvement of diagnosis and reporting of occupational cancers through medical consultations and the diffusion and access for health professionals, patients and the public to scientific and evidence-based information, with the creation of a website: <http://cancer-environnement.fr>. The Cancer and Environment Research Department is affiliated with the “Steroid signalling and breast tumor” team within the Cancer Research Centre of Lyon (CRCL, UMR Inserm 1052 CNRS 5286, Centre Léon Bérard).

In 2011, the Cancer and Environment Research Department, with the collaboration of the Observatoire Régional de Santé (ORS) in Rhône-Alpes, conducted a literature review on the impact of the management of household and similar waste (collection and sorting, composting, incineration and storage) on the health of waste management workers and the local population (Anzivino-Viricel et al. 2012). Results suggested a possible increase in breast cancer incidence in women living near the facilities, suggesting a potential association between breast cancer occurrence in women and exposure to MSWIs prior to the European standards. The authors recommended to further investigate areas

where scientific results are inconclusive, to improve exposure characterization and to take into account individual risk factors.

Moreover, the Institut national de la santé et de la recherche médicale (Inserm) and the French Agency for Food, Environmental and Occupational Health and Safety (ANSES) concluded in a report that research is needed about the impact of environmental exposure on breast cancer, and made some recommendations for future studies: to develop new methods to assess more accurately environmental exposure through GIS; to use epidemiological data from existing prospective cohorts and to further study gene-environment interactions (Inserm and Afsset 2008).

2.2.1. The GEO3N project

In this context, the Cancer and Environment Research Department, with the collaboration of the team “Lifestyle, genes and health”, CESP, Inserm U1018, Paris-Sud University, Institut Gustave Roussy, Villejuif, France, set up the GEO3N research project. GEO3N aims at investigating the hypothesis of an association between breast cancer occurrence and environmental exposures, with particular attention to dioxins that are emitted by incinerators, among women from the French E3N prospective cohort. The project proposed to overcome the limitations of existing studies with:

- The E3N prospective cohort, which collects regularly accurate information on nearly 100,000 women recruited throughout the national territory: residential history over a 19-year study period, main cancer risk factors (hormonal, reproductive and lifestyle-related factors), hormonal receptor status of breast cancers and dietary data from two diet history questionnaires;
- The assessment of airborne dioxin exposure based on a multi-source approach, considering several types of dioxin emitting sources (MSWI, metal and cement industries) and the evolution over time of industrial facilities’ technical characteristics;
- The implementation of a standardized methodology for the estimation of dioxin emission data of the facilities, using the Toolkit for Identification and Quantification of Releases of Dioxins, Furans and Other Unintentional POPs of the United Nations Environment Program (UNEP, (UNEP 2013)), in the absence of emission measurements for many facilities and periods;
- The development of a GIS-based metric integrating meteorological and topographical parameters, to assess airborne dioxin exposure at the individual address level over a long time period;
- The assessment of the main route of dioxin exposure in humans: the dietary dioxin exposure, by combining data from an E3N diet history questionnaire (DHQ) with data on food contamination by dioxins;
- Biological samples (blood or saliva) available for exploring gene-environment interactions.

Women of the E3N cohort are members of the Mutuelle Générale de l’Éducation Nationale (MGEN), a health insurance plan covering mostly teachers and workers of the Education sector. Therefore, occupational exposure to dioxins of the study subjects is considered to be negligible.

A first study was conducted in the Rhône-Alpes region to develop the methodology for the assessment of airborne dioxin exposure. This region was chosen because of its dense industrial network, making it an ideal geographic area with a relatively easy access to data.

Study design

The GEO3N research project relies on the E3N prospective cohort, and includes all the incident breast cancer cases identified in the cohort between 1990 and 2008 and validated with pathology reports (N=5,455). Because the assessment of airborne dioxin exposure required the geocoding of participants’ addresses on the whole national territory and over a 19-year period, a task that is time and

resource consuming, we designed a case-control study nested within the E3N prospective cohort to assess airborne dioxin exposure and estimate its association with breast cancer. This study design, along with an incidence density sampling, reduces the number of subjects for which information on exposure must be collected, while remaining within the context of a prospective cohort (Mantel 1973). Cases were matched to randomly selected controls (ratio 1:1) on age, date and department of residence at baseline, menopausal status and existence of a biological sample (blood, saliva, none). The study conducted in the Rhône-Alpes region was restricted to 429 cases and 786 controls (ratios 1:1 and 1:2) that had permanently resided in this region over the study period.

Chapters III and IV on the relationships between airborne dioxin exposure and breast cancer risk were investigated in the GEO3N nested case-control study; chapter II on the relationship between dietary dioxin exposure and breast cancer risk was investigated among all E3N women that fulfilled the 1993 DHQ (N=63,830).

The project team

The project involves a pluri-disciplinary collaboration with expertise in oncology, epidemiology, geography, environmental science and biology. The team members are: Dr. Béatrice Fervers (BF, M.D., HDR, associate professor), Laure Dossus (LD, PhD, epidemiologist, 2011-2015), Elodie Faure (EF, geomatician, 2012-now), Lucie Anzivino-Viricel (LAV, PhD, environmental epidemiologist, 2012-2013), Aurélie Danjou (AD, PhD student in epidemiology, 2013-now), Thomas Coudon (TC, PhD student in environmental science, 2014-now), Guillaume Harel and Camille Denis (GH, 2014-2015, and CD, 2015-2016, environmental engineers), Charlotte Carretero (CCa, geocoding technician, 2015-2016), Delphine Praud (DP, Post-Doc, epidemiologist, 2016-now), Maxime Guillou (MG, geocoding technician, 2016-now) and Xavier Morelli (XM, Post-Doc, epidemiologist-SIG, 2016-now).

The project partners are:

- Françoise Clavel-Chapelon (FCC, DR1, PhD, epidemiologist) from the team Team “Lifestyle, genes and health”, Inserm U1018, Center for Research in Epidemiology and Population Health (CESP), Université Paris-Sud;
- Pietro Salizzoni (PS, MCU, environmental engineering) and Chi-Vuong Nguyen (CVN, PhD student in environmental science), from the Fluid Mechanics and Acoustics Laboratory, UMR 5509, Ecole Centrale de Lyon;
- Karen Leffondré (KL, MCF, HDR) and Emilie Lévêque (EL, PhD student in biostatistics), from the team “Epidemiology and Biostatistics”, Inserm U1219, Institut de Santé Publique, d’Épidémiologie et de Développement (ISPED), Université de Bordeaux.

A scientific committee has been constituted to advise the team in the development of the project methodology. It includes the Santé Publique France Agency (formerly Institut de Veille Sanitaire), the French Geological Survey (BRGM), the National competence center for Industrial Safety and Environmental Protection (INERIS), the Interprofessional Technical Centre for Studies on Air Pollution (CITEPA), the French Environment and Energy Management Agency (ADEME), the Direction Régionale de l’Environnement, de l’Aménagement et du Logement Auvergne-Rhône-Alpes (DREAL RA), ANSES, the Association Agréée Surveillance Qualité de l’Air (Air Rhône-Alpes) and INSAValor.

Funding

Project funding includes national grant from ADEME and regional grant from the Ligue Contre le Cancer 71. The study conducted in the Rhône-Alpes region was funded by the Oncostarter Program of

the Cancéropôle Lyon Auvergne Rhône-Alpes (CLARA) and the Région Auvergne Rhône-Alpes. Two PhD students (Thomas Coudon and I) received a 3-year academic doctoral fellowship from EDISS, Université de Lyon. I also received a mobility grant from the CLARA for training, within the team Epidemiology and Biostatistics (ISPED, Université de Bordeaux), on specific statistical methods that were applied in the project.

2.2.2. The E3N cohort study

The E3N cohort is an ongoing prospective study aiming at identifying environmental causes of major chronic diseases, and particularly cancer, in women (Clavel-Chapelon et al. 1997; Clavel-Chapelon and for the E3N Study Group 2015; Inserm U1018 2016). The cohort is conducted within the team “Lifestyle, genes and health” (CESP, Inserm U1018) at the Institut Gustave Roussy, and was first led by Françoise Clavel-Chapelon, DR1, then since 2015, by Marie-Christine Boutron-Ruault, DR2.

E3N was initiated in 1990, when nearly 100,000 women, born between 1925 and 1950, and insured by the MGEN, a national health insurance plan primarily covering teachers, volunteered and completed a baseline self-administered questionnaire. The participants also signed an informed consent form that allowed collecting additional information on their vital status, changes of address and medical fee reimbursements from the MGEN.

Since baseline, the participants filled in questionnaires every 2 to 3 years about lifestyle-related and metabolic factors, including nutrition, physical activity, anthropometric characteristics, hormonal treatment use and reproductive factors; medical and familial histories; and health status. To date, eleven questionnaires have been sent; the response rate remained stable at approximately 80% at each questionnaire. From 1994 to 1999, blood samples were collected among nearly 25,000 women; the biological material bank was then complemented in 2009-2011 with saliva samples from about 47,000 women who had not provided blood sample.

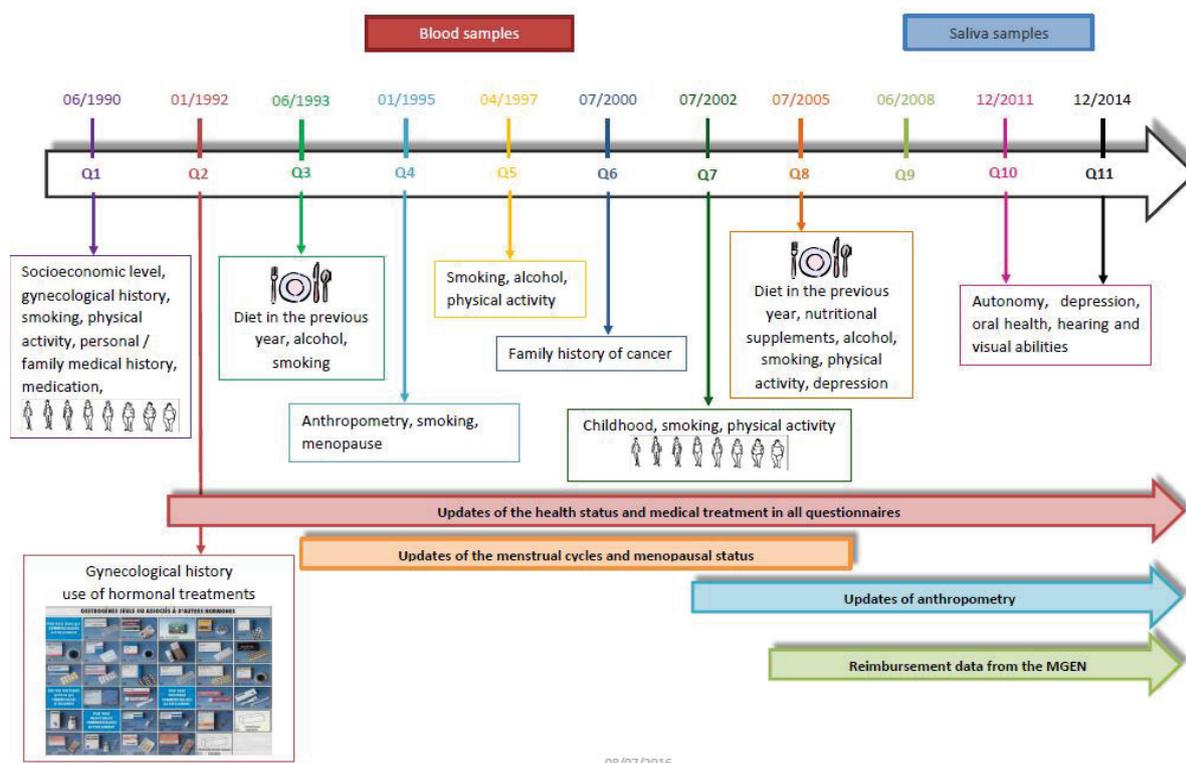


Figure 6. Calendar of the E3N self-administered questionnaires, 1990-2014 (last update on 07-2016).

The E3N study protocol was approved by the French National Commission for Computed Data and Individual Freedom (CNIL, Commission Nationale Informatique et Libertés). The Inserm, the MGEN, the Ligue contre le cancer and the Institut Gustave Roussy are the founding partners and support the implementation of the E3N cohort study.

Since 1993, E3N is the French component of EPIC, which aimed at investigating the links between diet, lifestyle, environmental factors and the occurrence of cancer and other chronic diseases based on a cohort of more than 500,000 men and women from ten European countries (Denmark, France, Germany, Greece, The Netherlands, Italy, Norway, Spain, Sweden and the United Kingdom).

Dietary data

Along with follow-up questionnaires, two diet history questionnaires have been sent to the participants in 1993 (DHQ3) and 2005 (DHQ8). The DHQ3 was composed of a qualitative section and a quantitative section, and a photo-booklet was attached to the DHQ3 to assist the participants in estimating portion sizes. In the quantitative section, questions covered 8 meals and snacks of the day (breakfast, morning snack, aperitif before lunch, lunch, snack in the afternoon, aperitif before dinner, dinner and snack after dinner). For each food item, food group and beverage, participants were asked to indicate their consumption frequency (per week or per month) and to estimate portion sizes using the photo-booklet (Figure 7).

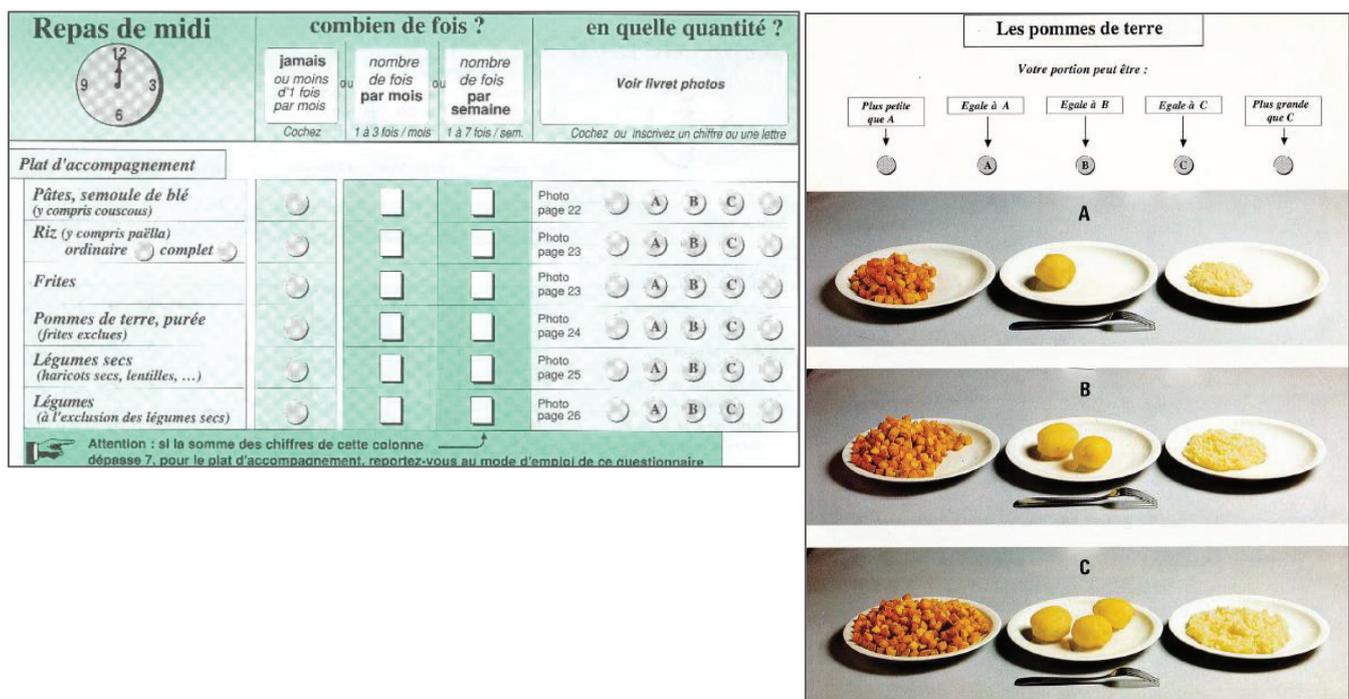


Figure 7. Extracts from the diet history questionnaire sent in 1993 (quantitative section, page 6) and the photo-booklet (page 24). E3N study, France

The qualitative section asked further questions about specific food items within each food group (Figure 8). Women could indicate the consumption frequency of each item (never or rarely “0”, sometimes “+”, regularly “++” and very often “+++”). The qualitative section also contained detailed questions on fat, sugar and cooking methods.

6. Les légumes chauds : En consommez-vous habituellement ?

Non, jamais (ou moins d'1 fois par mois), passez à la question 7.

Oui, si oui, lesquels ? (voir page 15 comment remplir le tableau)

	0	+	++	+++
Chou				
Chou-fleur				
Choux de Bruxelles				
Epinards				
Endives				
Champignons				
Céleri en branche				
Bettes (ou cardons)				
Fenouil cuit				

	0	+	++	+++
Haricots verts				
Carottes cuites				
Salsifis				
Tomates cuites				
Aubergines				
Poivrons cuits				
Courgettes				
Petits pois				
Autres (précisez) :				

 **Attention : pour chaque légume vous ne devez cocher qu'une seule case par ligne**

Figure 8. Extract from the diet history questionnaire sent in 1993 (qualitative section, page 17). E3N study, France

Each food portion illustrated in the photo-booklet reflected a quantity in g or mL (then converted to g). This amount was multiplied by the food consumption frequency in order to obtain an amount per week or per month that was then converted to g per day. The DHQ3 provided dietary information for 208 food items.

A validation study was performed in 1995 to evaluate the “quality” of the photo-booklet for the assessment of food portion sizes, and test whether quantities and volumes were carefully illustrated by the photographs and whether the range of portions selected included the full range of food quantities actually consumed (Lucas et al. 1995). Moreover, another validation study that was performed on the DHQ3 in 1997 showed that the questionnaire had an overall good reproducibility and could be used to classify the study subjects according to their food consumption over a one-year period (Van Liere et al. 1997).

Residential history data

Women’s residential history was built through the follow-up questionnaires, *i.e.* from 1990 to present. The home addresses from the participants were recorded for the first and fifth to eleventh questionnaires; only zip code were recorded from the third and fourth questionnaires; there was no address recorded in the database for the second questionnaire. The place of birth (zip code and commune) was obtained from the first questionnaire.

CHAPTER II. DIETARY DIOXIN EXPOSURE AND BREAST CANCER RISK

Assessment of dietary dioxin exposure and estimation of breast cancer risk

1. Introduction

In the absence of occupational and accidental exposures, humans are exposed to low levels of dioxins through food, water, soil and atmosphere. Dioxins contaminate food chains and accumulate at high trophic levels, including livestock and humans (Van den Berg et al. 1994). Because of the lipophilic nature of dioxins, liver and adipose tissue are their main storage sites in the body, although dioxins can be detected in serum and mother's breast milk. The general population is primarily concerned with dietary dioxin exposure, and dairy products, seafood, meat and eggs – particularly rich in fat – are the most contaminated feed (Fries 1995; Parzefall 2002; IARC 2012c).

Levels of dioxins depend on the diet composition, the individual's metabolism and several parameters that influence excretion of dioxins such as breastfeeding and body weight changes (Wolff et al. 2005). In a French study investigating the dioxin impregnation in a population living in the vicinity of a MSWI, levels of dioxins were associated with age, sex, body shape, recent weight change, smoking habits, current socio-professional status and food (Fréry et al. 2009). A high consumption of fish was associated with high serum concentration of dioxins in several studies (Arisawa et al. 2003, 2011; Chen et al. 2007; Svensson et al. 1991; Turunen et al. 2010; Uemura et al. 2008); while breastfeeding led to lower serum concentrations (Arisawa et al. 2003). The dioxin concentration in breast milk was found lower among primiparous smoking mothers, and higher among fish eater mothers (Takekuma et al. 2004).

The dietary dioxin exposure in the general population has been previously assessed by combining food contamination data by dioxins with food consumption data from dietary questionnaires, and specific food groups contributing to the exposure have been highlighted. Several studies reported that the major contributors to dietary dioxin intake were dairy products, fat, meat and seafood (Bilau et al. 2008; De Mul et al. 2008). However, contribution of specific food groups to the dietary intake depends on the consumption of the latter and other studies found that cereal products and fruits were also major contributors (Domingo et al. 1999).

In France, national monitoring programs have been implemented since 1996 to analyze PCDD and PCDF contamination in food products consumed by the general population. The dietary dioxin intake was also assessed using food consumption data from French national dietary surveys. So far, three assessments of the dioxin dietary intake of the French population have been conducted in 1996-98, 2001-04 and 2007-08 (CSHPF 2000; Sirot et al. 2009; Tard et al. 2007).

Although dioxins are suspected to increase breast cancer risk and dietary intake is considered to constitute the main route of environmental dioxin exposure (Hattemer-Frey and Travis 1989; IARC 2012c), to date, no epidemiological study has investigated the association between dietary dioxin exposure and risk of breast cancer.

Objectives and methods

The objectives of this chapter were to assess dietary dioxin exposure of the E3N women and to estimate the risk of breast cancer associated with this exposure. The exposure assessment and risk analyses were performed within the E3N cohort, among the 63,830 women that completed the DHQ3 in 1993 and were followed until breast cancer diagnosis, death or return of the 2008 questionnaire, whichever occurred first. A total of 3,465 women reported primary breast cancer between 1993 and 2008. End of follow-up was set to 2008 (ninth E3N questionnaire) because self-reported breast cancer cases were still being validated beyond that date at the time of analysis.

To assess the individual dietary dioxin exposure of the E3N women, we used:

- Food consumption data from the validated DHQ3 of 1993 that reported daily quantity (in g per

day) for 208 food and beverage items;

- Food contamination data by dioxins, reported by the French High Council on Public Health in 2000 (CSHPF 2000). Dioxin concentrations have been measured in 444 food samples collected between 1996 and 1998 from several control and monitoring plans, performed by the Directorate for food of the French Agriculture ministry (DGAL), the General Directorate for Competition Policy, Consumer Affairs and Fraud Control (DGCCRF) and the Agrifood profession (Article #1). These contamination data were the closest to the period during which the DHQ3 was administered;
- An assessment method recommended by WHO that consisted in combining food consumption data with food contamination data, considering the fat content of foods and women's body weight (FAO/WHO 2005).

The association between dietary dioxin exposure and breast cancer risk was estimated using Cox's proportional hazard models and stratified by birth cohort. Breast cancer risks were estimated for an increase of 1 standard deviation and by quartiles of dietary dioxin exposure. Models were adjusted for established breast cancer risk factors and potential confounders; some factors were considered as time-dependent variables. Sub-group analyses were performed according to factors that could influence the association: menopausal status and BMI (Tehard et al. 2004); weight change during follow-up; dietary pattern (Cottet et al. 2009); breastfeeding and hormone receptor status (ER/PR). Due to the ability of flavonoids to protect against the toxicity of dioxins (Ashida et al. 2000), additional adjustment on the flavonoid consumption was advised.

2. Article #1

Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort

Aurélie MN Danjou^{1,2}, Béatrice Fervers^{1,2}, Marie-Christine Boutron-Ruault^{3,4,5}, Thierry Philip^{1,2}, Françoise Clavel-Chapelon^{3,4,5} and Laure Dossus^{1,3,4}

¹ Unité Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France

² Université Claude Bernard Lyon 1, 43 boulevard du 11 Novembre 1918, 69100 Villeurbanne, France

³ Inserm, Center for Research in Epidemiology and Population Health (CESP), U1018, Nutrition, Hormones and Women's Health team, 114 rue Edouard –Vaillant, 94805 Villejuif Cedex, France

⁴ Université Paris-Sud, UMRS 1018, Villejuif, France

⁵ Gustave Roussy, F-94805 Villejuif, France

Reference:

Aurélie Danjou, Béatrice Fervers, Marie-Christine Boutron-Ruault, Thierry Philip, Françoise Clavel-Chapelon, and Laure Dossus. 2015. 'Estimated Dietary Dioxin Exposure and Breast Cancer Risk among Women from the French E3N Prospective Cohort'. *Breast Cancer Research* 17 (1): 39. doi:10.1186/s13058-015-0536-9.

RESEARCH ARTICLE

Open Access

Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort

Aurélien MN Danjou^{1,2}, Béatrice Fervers^{1,2*}, Marie-Christine Boutron-Ruault^{3,4,5}, Thierry Philip^{1,2},
Françoise Clavel-Chapelon^{3,4,5} and Laure Dossus^{1,3,4}

Abstract

Introduction: Dioxins are environmental and persistent pollutants mostly emitted from combustion facilities (e.g. waste incinerators, metal and cement industries). Known to be endocrine disrupting chemicals, dioxins are suspected to increase breast cancer (BC) risk. Although diet is considered the primary source of dioxin exposure, no previous study has been published on dietary dioxin exposure in relation to BC risk. We aimed to assess dietary dioxin exposure among women from the E3N cohort and estimate BC risk associated with this exposure.

Methods: The study included 63,830 women from the E3N cohort who completed a diet history questionnaire (DHQ) in 1993 and were followed until 2008. Dietary dioxin exposure was estimated by combining consumption data from the E3N DHQ and food dioxin contamination data from a French national monitoring program. Hazard ratios (HR) and 95% confidence intervals (CI) were estimated by Cox models adjusted for BC risk factors.

Results: Mean dietary dioxin exposure was estimated at 1.3 ± 0.4 pg/kg body weight (BW)/day. A 0.4 pg/kg BW/d increase in dioxin intake was not associated with overall BC risk (HR = 1.00; 95% CI: 0.96, 1.05). A significant decrease in risk of estrogen receptor negative (ER-)/progesterone receptor negative (PR-) tumors was observed among post-menopausal women in the upper quartile of estimated dioxin intake (HR for Q4 vs. Q1: 0.65; 95% CI: 0.45, 0.96; *P* for trend across quartiles = 0.0463).

Conclusions: Overall, no association between estimated dietary dioxin exposure and BC risk was found among E3N women. Further studies should include both dietary and environmental exposures to determine whether low-dose dioxin exposure is associated with BC risk.

Introduction

Polychlorinated dibenzo-*p*-dioxins (PCDD) and polychlorinated dibenzofurans (PCDF), together defined as dioxins, are toxic environmental by-products resulting from the incomplete combustion of chlorinated substances from metal and cement industries, waste incinerators or domestic activities. Because they are highly lipophilic and persistent pollutants, dioxins can contaminate the food chain and accumulate in adipose tissues [1]. Therefore, in the general population, diet is considered

one of the main sources of exposure to dioxins [1-3]. Most of the studies assessing dietary dioxin exposure concluded that food from animal origin rich in fat such as dairy products and meat, and seafood are the most contaminated foods. Because contribution of specific food groups to dioxin exposure depends on their consumption, main contributors might differ between populations due to different dietary patterns.

Dioxins are known to be endocrine disrupting chemicals [4] and suspected to play a role in the increase of hormone-related cancer incidence such as breast cancer. The 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) is the most toxic congener and has been classified as carcinogenic to humans by the International Agency for Research on Cancer (IARC) and the US Environmental Protection Agency (US EPA) [5,6].

* Correspondence: beatrice.fervers@lyon.unicancer.fr

¹Unité Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon, Cedex 08, France

²Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69100 Villeurbanne, France

Full list of author information is available at the end of the article

Few epidemiological studies have investigated breast cancer risk associated with dioxin exposure, and there have been inconsistent results. Among the cohort of residents exposed to TCDD following the Seveso accident (Italy, 1976), no increased breast cancer risk was observed among women living in the areas most exposed to TCDD [7,8]. Nevertheless, elevated serum TCDD levels measured in Seveso women who lived in the two most contaminated zones were associated with increased breast cancer risk in a first analysis conducted in 2002 [9]. However, this was not confirmed with longer follow up of the same population [10]. Studies conducted among cohorts of workers exposed to dioxins during production of herbicides showed higher breast cancer mortality in these populations (reviewed in [11,12]). Other studies also reported elevated breast cancer mortality [13] and increased breast cancer risk [14] in cohorts of community residents exposed to dioxins emitted by industrial facilities. However in a French case-control study, a decrease in breast cancer risk was observed among older women living near a municipal solid-waste incinerator (MSWI) and most exposed to dioxins [15].

To our knowledge no epidemiological study has explored the association between dioxin exposure through diet and breast cancer risk. Therefore, our study aimed to assess dietary dioxin exposure among women from the French E3N prospective cohort and estimate breast cancer risk associated with this exposure.

Methods

The E3N cohort study

The E3N (*Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale*) study is an ongoing prospective cohort involving 98,995 French female volunteers, born between 1925 and 1950 and members of a national teachers' health insurance plan [16]. The E3N study was initiated to identify female cancer risk factors, especially dietary, hormonal and reproductive factors. The study began in 1990 when participants returned the first questionnaire and gave signed informed consent. Since then women have completed self-administered questionnaires sent by mail every 2 to 3 years about their lifestyle, health status and medical history. The E3N study protocol was approved by the French National Commission for Data Protection and Privacy.

Study population

Women included in our study population were participants of the E3N study, who had completed the E3N diet history questionnaire (DHQ) in 1993 and were followed until June 2008. Among the 74,522 women with dietary data, women having under- or over-reported their consumption, that is, those in the bottom and top 1% of the energy intake to

basal metabolic rate ratio computed on the basis of age, height and weight, were excluded ($n = 1,360$). In addition, we excluded women who reported cancer diagnosis (except non-melanoma skin cancer) before completing the E3N DHQ ($n = 4,705$) along with those who developed *in situ* breast cancer ($n = 483$) and those for whom follow-up information was unavailable ($n = 822$). We further excluded 3,302 women with missing body weight (BW) in 1993 and 20 women who had never menstruated. Finally 63,830 women were included in our study population.

Breast cancer cases

From the third self-administered questionnaire to the ninth, data on health status and medical history of women, including cancer occurrence, were collected and updated. A total of 3,465 women reported primary invasive breast cancer between 1993 and 2008 and 92% of them were confirmed by pathology reports. Information on estrogen receptor (ER) and progesterone receptor (PR) were extracted from these reports, and invasive breast cancer cases were classified as ER+/PR+ ($n = 1,596$), ER+/PR- ($n = 561$) or ER-/PR- ($n = 414$). Because of the small number of cases, the ER-/PR+ invasive breast cancers were not considered in the analyses ($n = 107$). Because the proportion of false-positive self-reported breast cancer in the E3N cohort was <5%, breast cancers not confirmed by pathology report ($n = 285$) were considered as cases. Deaths were identified through family contact or from insurance files, and causes of death were obtained from the French National Service on Causes of Death [17].

Assessment of dietary dioxin exposure

Consumption data

Dietary data were collected from a validated self-administered DHQ, specifically developed for the assessment of the previous year's usual diet of the E3N population and sent in 1993 [18]. The E3N DHQ covered the daily consumption of 208 food items by collecting food frequencies and portion sizes for 8 meals and snacks during the day. The questionnaire comprised both a quantitative and a qualitative part. In the quantitative part, women were asked to indicate the frequency of consumption and to estimate portion sizes of food groups, food items and beverages. Portion sizes were estimated with a validated photo booklet [19]. The qualitative part provided detailed information on women's relative consumption frequency of specific food items within one of the food groups of the first part. Furthermore, a food composition table was generated from a French national database [20], providing information on fat content of most of the food items assessed in the E3N DHQ.

Contamination data

An assessment of the level of food contamination by dioxins (PCDDs and PCDFs) in France was published in 2000 by the French High Council for Public Health (*Conseil Supérieur d'Hygiène Publique de France* (CSHPF)) [21]. These contamination data were based on food samples collected from 1996 to 1998 which was the closest period to the E3N DHQ. Food samples analyzed in this report ($n = 444$) included dairy products, meats, seafood products, cereal products, fruits and vegetables, eggs and fats (Additional file 1: Table S1). They were collected from monitoring and control plans mostly performed by the Directorate for food of the French Agriculture ministry (DGAL) and the General Directorate for Competition Policy, Consumer Affairs and Fraud Control (DGCCRF). Measurements of dioxin contamination in these samples were performed in one single laboratory (Carso Laboratory, Lyon, France) with a certified method using high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS). Dioxin concentrations were calculated with the lower-bound method, assuming that non-detected congeners were equal to zero. Results were expressed in pg TEQ/g lipid weight, except for foods from vegetal origin (that is, cereal products and fruit and vegetables) for which results were expressed in pg TEQ/g fresh weight.

Exposure assessment

Dietary dioxin exposure in our study was estimated for each woman, by combining consumption data from the E3N DHQ and lipid levels from the food composition table developed by the E3N team with food dioxins contamination data from the CSHPF report (Additional file 1: Table S1), using the following formulas recommended by the World Health Organization (WHO) [22]:

Equation 1:

$$E_i = \frac{\sum_{k=1}^n C_{i,k} \times (F_k \times L_k)}{BW_i} \quad (1)$$

Equation 2, for cereal products and fruit and vegetables:

$$E_i = \frac{\sum_{k=1}^n C_{i,k} \times P_k}{BW_i} \quad (2)$$

where E_i is the dietary dioxin exposure for the woman i (in pg TEQ/kg BW/day), $C_{i,k}$ is the consumption of the food k by the woman i (in g/day), F_k is the proportion of lipids of the food k (in %), L_k is the average dioxin contamination of the food k (in pg TEQ/g lipid weight, see equation 1), P_k is the average dioxin contamination of the food k (in pg TEQ/g fresh weight, see equation 2) and BW_i is the body weight of the woman i (in kg) reported in the E3N questionnaire of 1993.

The intake of PCDDs and PCDFs was calculated in toxic equivalent (TEQ) using toxic equivalence factors (TEF) first developed by the North Atlantic Treaty Organization (NATO) in 1989 [23]. In 2001, the WHO established a dioxin-tolerable daily intake of 2.3 pg TEQ/kg BW/day [24,25].

Statistical analysis

Participants contributed person-years of follow up from the date they completed the E3N DHQ to the date of diagnosis of primary invasive breast cancer, date of death, date of return of the ninth questionnaire or June 2008, whichever occurred first. Estimated dietary dioxin exposure was categorized according to quartiles of its distribution in the study population. Baseline characteristics of the participants were described by quartiles of estimated dietary dioxin exposure using mean and standard deviation (SD) for quantitative variables and frequency and percentage for qualitative variables.

Hazard ratios (HR) and 95% confidence intervals (CI) for breast cancer were estimated for an increase of 0.43 pg TEQ/kg BW/day (that is, 1 SD) and for quartiles of estimated dietary dioxin exposure with first quartile used as reference, using Cox proportional hazard models, stratified by 5-year-interval birth cohorts (from 1925 to 1950). Women's age was used as the timescale. To evaluate potential dose-response associations, tests for linear trend across quartiles were derived from the Wald test of the models with the semi-continuous variable. All models were adjusted for known breast cancer risk factors and potential confounders: height (cm), body mass index (BMI) before and after menopause (cut off: 25, 30 kg/m²; considered as a time-dependent variable), total energy intake excluding alcohol (kcal/day), alcohol intake (cut off: 0, 6.9 g/day), education (undergraduate/postgraduate with a 1- to 4-year university degree/postgraduate with a 5+ year university degree), physical activity at baseline (cut off: 34, 46, 62 metabolic equivalent task-hour per week (MET-h/w)), smoking status (never smoked/ex-smoker/current smoker; time-dependent variable), age at menarche (cut off: 12, 14 years), previous use of oral contraceptives (never/ever), previous use of progestin before menopause (never/ever), menopausal status combined with use of menopausal hormone treatment (MHT) (premenopausal/postmenopausal non-using of MHT/postmenopausal using of MHT; time-dependent variable), age at menopause (cut off: 47, 54 years, among postmenopausal women only), age at first full-term pregnancy combined with number of live births (no child/ ≤ 2 children and age at first birth < 30 years/ ≤ 2 children and age at first birth ≥ 30 years/ ≥ 3 children), breastfeeding (never/ever), previous family history of breast cancer (yes/no), previous history of personal benign breast disease (yes/no; time-dependent variable), previous

mammography (yes/no; time-dependent variable). Further adjustments for food group consumption (including fruit and vegetable, fish and seafood, dairy products or total dietary flavonoid intake) were made. All adjustment variables had less than 5% missing data that were replaced by their modal or median value in the population. Information on estimated dietary dioxin exposure was available for all women of the study population.

We also estimated risk of invasive breast cancer according to hormone receptor status (ER/PR). We excluded cases for whom joint hormone-receptor status (n = 787) was missing from the corresponding analyses. Interactions between BMI and estimated dietary dioxin exposure were tested among premenopausal and postmenopausal women in relation to breast cancer risk because of opposite associations between BMI and risk of breast cancer before and after menopause [26] and the storage of dioxins in adipose tissues. Analyses were also stratified by weight change during follow up (cut points >-2 kg/5 years; >2 kg/5 years) and dietary patterns ("Western" vs. "Healthy" patterns [27]). Because of previously observed negative associations, we studied interactions between estimated dietary dioxin exposure and smoking status, and breastfeeding [28,29] while interaction between estimated dietary dioxin exposure and alcohol consumption was studied due to potential positive association [30]. All tests for interaction were derived from the likelihood-ratio test comparing the models with and without an interaction term. Sensitivity analysis was performed, excluding breast cancer cases diagnosed less than 2 years after completing the E3N DHQ, to eliminate possible preclinical tumors. All P-values were two-sided and the significance level was set at 0.05. We used the SAS statistical software version 9.3 (SAS Institute Inc., Cary, NC, USA) for data analysis.

Results

We analyzed data from 63,830 women with a median follow up of 14.9 years, corresponding to 888,505 person-

years and during which 3,465 incident invasive breast cancers occurred. The food groups most consumed were fruit and vegetables (mean ± SD: 724.8 ± 288.5 g/d), dairy products (341.0 ± 200.1 g/d) and cereal products (247.7 ± 108.8 g/d, Table 1). Fish and seafood were less frequently consumed (38.0 ± 27.2 g/d). The average level of dietary dioxin exposure was estimated at 1.3 ± 0.4 pg TEQ/kg BW/d (range: 0.1 to 5.7 pg TEQ/kg BW/d) and 2.7% of women exceeded the 2.3 pg TEQ/kg BW/d toxicity threshold established by the WHO [24,25]. The highest contributors to the total estimated dietary dioxin exposure were dairy products (33.3%), fruit and vegetables (22.1%), meat (18.8%) and fish and seafood (15.6%), while the contribution from eggs (5.5%), cereal products (3.6%) and added fats (1.2%) was low.

Higher estimated dietary dioxin exposure was seen in younger women and in those with higher energy intake, higher alcohol consumption, lower pre- and postmenopausal BMI, and in women with 1 or 2 children before 30 years, with a personal history of benign breast disease and those who had breastfed at least one of their children (Table 2).

Table 3 shows HRs associated with estimated dietary dioxin exposure among overall participants and according to hormone receptor status (ER/PR). An increase of 0.43 pg TEQ/kg BW/d in the intake of dioxins was not associated with risk of invasive breast cancer risk among overall participants, neither in the univariable nor in the multivariable models: HR = 1.01 (0.99, 1.03) and HR = 1.00 (0.96, 1.05), respectively. There was no statistically significant linear trend between quartiles of estimated dietary dioxin exposure and breast cancer risk (P for trend = 0.9405). There was also no association between estimated dietary dioxin exposure and ER+/PR+ or ER+/PR- breast cancer risk. However a borderline statistically significant decrease in ER-/PR- breast cancer risk was observed across quartiles of estimated dietary dioxin exposure (HR for Q1 vs. Q4: 0.72, 95% CI: 0.50, 1.02; P for trend across quartiles = 0.0629). When stratifying by menopausal status, we

Table 1 Contribution of food groups to estimated dietary dioxin exposure in the E3N cohort, 1993 (n = 63,830)

Food groups	Consumption (g/day), mean ± SD	Dietary dioxin exposure (pg/kg body weight/day), mean ± SD	Contribution to dietary dioxin exposure, %
Dairy products	341.0 ± 417.7	0.428 ± 0.223	33.3
Fruit and vegetables	724.8 ± 288.5	0.284 ± 0.116	22.1
Meat	122.2 ± 55.2	0.242 ± 0.184	18.8
Seafood	38.0 ± 27.2	0.200 ± 0.173	15.6
Eggs	26.4 ± 20.4	0.071 ± 0.054	5.5
Cereal products	247.7 ± 108.8	0.046 ± 0.020	3.6
Added fats	25.2 ± 11.7	0.015 ± 0.013	1.2
Total	1,525.3 ± 417.7	1.285 ± 0.428	100.0

Table 2 Main baseline characteristics of the study population according to quartiles of estimated dietary dioxin exposure, E3N cohort, 1993 (n = 63,830)

	Dietary dioxin exposure (pg/kg body weight/day)			
	<0.98	[0.98, 1.23]	[1.23, 1.52]	≥1.52
Age (years), mean ± SD	53.5 ± 6.8	53.0 ± 6.7	52.5 ± 6.5	51.9 ± 6.3
Energy intake (kcal/day), mean ± SD	1710.3 ± 391.7	2003.6 ± 405.5	2231.6 ± 437.0	2580.8 ± 513.4
Smoking status, n (%)				
Non-smoker	8672 (54.3)	8611 (54.0)	8614 (54.0)	8429 (52.8)
Former smoker	5102 (32.0)	5259 (33.0)	5243 (32.9)	5288 (33.1)
Smoker	2184 (13.7)	2087 (13.1)	2101 (13.2)	2240 (14.0)
Combined menopausal status, MHT^a use and BMI^b, n (%)				
Premenopausal				
BMI <25 kg/m ²	4797 (73.7)	5665 (82.2)	6362 (87.1)	7143 (91.5)
BMI ≥25 kg/m ²	1715 (26.3)	1227 (17.8)	942 (12.9)	667 (8.5)
Postmenopausal not using MHT				
BMI <25 kg/m ²	2572 (58.7)	2679 (70.4)	2715 (76.5)	2692 (82.8)
BMI ≥25 kg/m ²	1813 (41.4)	1129 (29.7)	833 (23.5)	560 (17.2)
Postmenopausal using MHT				
BMI <25 kg/m ²	3526 (69.7)	4127 (78.5)	7261 (83.5)	4344 (88.7)
BMI ≥25 kg/m ²	1535 (30.3)	1130 (21.5)	845 (16.6)	551 (11.3)
Parity and age at first birth, n (%)				
No child	2035 (12.8)	1890 (11.8)	1793 (11.2)	1685 (10.6)
≤2 children <30 years old	7661 (48.0)	7913 (49.6)	8076 (50.6)	8393 (52.6)
≤2 children ≥30 years old	1527 (9.6)	1402 (8.8)	1464 (9.2)	1378 (8.6)
≥3 children	4735 (29.7)	4752 (29.8)	4625 (29.0)	4501 (28.2)
Breastfeeding, n (%)				
Ever	8916 (55.9)	9234 (55.9)	9321 (58.4)	9365 (58.7)
Never	7042 (44.1)	6723 (42.1)	6637 (41.6)	6592 (41.3)
Alcohol intake, n (%)				
0 g/day	2636 (16.5)	2021 (12.7)	1685 (10.6)	1356 (8.5)
<6.9 g/day	6688 (41.9)	6354 (39.8)	5897 (37.0)	5181 (32.5)
≥6.9 g/day	6634 (41.6)	7582 (47.5)	8376 (52.5)	9420 (59.0)
Personal history of benign breast disease, n (%)				
Yes	4292 (26.9)	4570 (28.6)	4907 (30.8)	4949 (31.0)
No	11666 (73.1)	11387 (71.4)	11051 (69.3)	11008 (69.0)

^aMHT, menopausal hormone treatment; ^bBMI, body mass index.

observed a statistically significant decrease in ER-/PR-breast cancer risk (HR for Q4 vs. Q1: 0.65, 95% CI: 0.45, 0.96; *P* for trend across quartiles = 0.0463) (Table 4).

Adjusting for fruit and vegetable consumption, total dietary flavonoid intake, or any other food groups, did not change the results (data not shown). There was no significant interaction between BMI and estimated dietary dioxin exposure among pre- and postmenopausal women in relation to invasive breast cancer risk. There was no effect modification by smoking, alcohol consumption, dietary patterns, gain or loss of weight during follow up and breastfeeding, and there was no association

between estimated dietary dioxin exposure and breast cancer risk in any above-defined subgroups (Additional file 1: Table S2). Our results did not change when excluding breast cancer cases diagnosed less than 2 years after inclusion (data not shown).

Discussion

To our knowledge, this study is the first to assess breast cancer risk associated with dietary dioxin exposure. Among women from the E3N prospective cohort, we did not observe any increase of breast cancer risk associated with estimated dietary dioxin exposure. Among

Table 3 Hazard ratios for invasive breast cancer for increased intake of 0.43 pg/kg body weight/day, according to quartiles of estimated dietary dioxin exposure and hormone receptor status (n = 63,830), 1993 to 2008

Populations ^c	/0.43	Dietary dioxin exposure (pg/kg body weight/day)				P-trend
		<0.98 (ref)	[0.98, 1.23]	[1.23, 1.52]	≥1.52	
Cases, n	3465	880	853	848	884	
HR (95% CI) ^a	1.01 (0.99, 1.03)	1.00	0.97 (0.88, 1.06)	0.96 (0.88, 1.06)	1.02 (0.92, 1.11)	0.4181
HR (95% CI) ^b	1.00 (0.96, 1.05)	1.00	0.94 (0.86, 1.04)	0.93 (0.83, 1.03)	0.96 (0.85, 1.09)	0.9405
ER+/PR+^d						
Cases, n	1596	410	388	389	409	
HR (95% CI) ^a	1.02 (0.98, 1.07)	1.00	0.94 (0.82, 1.08)	0.95 (0.82, 1.09)	1.00 (0.88, 1.15)	0.5328
HR (95% CI) ^b	1.00 (0.93, 1.07)	1.00	0.91 (0.79, 1.06)	0.90 (0.77, 1.05)	0.94 (0.78, 1.12)	0.9678
ER+/PR-^d						
Cases, n	561	131	146	137	147	
HR (95% CI) ^a	1.04 (0.96, 1.13)	1.00	1.12 (0.88, 1.41)	1.05 (0.83, 1.34)	1.14 (0.90, 1.45)	0.3565
HR (95% CI) ^b	1.04 (0.93, 1.17)	1.00	1.10 (0.86, 1.41)	1.03 (0.79, 1.35)	1.14 (0.84, 1.54)	0.4571
ER-/PR-^d						
Cases, n	414	109	103	104	98	
HR (95% CI) ^a	0.96 (0.87, 1.06)	1.00	0.93 (0.71, 1.22)	0.94 (0.72, 1.23)	0.89 (0.67, 1.16)	0.4221
HR (95% CI) ^b	0.88 (0.77, 1.01)	1.00	0.86 (0.65, 1.13)	0.81 (0.60, 1.10)	0.72 (0.50, 1.02)	0.0629

^aAge-adjusted models; ^badjusted for age, height, body mass index, energy intake, education, physical activity, smoking status, menopausal status combined with use of menopausal hormone treatment, alcohol intake, age at menarche, use of oral contraceptives, use of progestin, age at menopause, age at first full-term pregnancy and number of live births, breastfeeding, family history of breast cancer, history of personal benign breast disease and mammography; ^cER-/PR+ invasive breast cancer risk was not analyzed due to small number of cases (n = 107); ^dn = 787 invasive breast cancer cases with missing hormone receptor status. HR, hazard ratio.

postmenopausal women, an inverse association was observed between estimated dioxin intake and ER-/PR-breast cancer risk.

Our estimation of an average dioxin intake of 1.3 pg TEQ/kg BW/d is of the same order of magnitude or lower than in other studies. In a monitoring program conducted in 1996 to 1998 in the general French population and using the same dioxin contamination data as in our study, mean dietary dioxin exposure was also estimated at 1.3 pg TEQ/kg BW/d [21]. At the same period in Europe, estimated dietary dioxin exposure of the general population varied between 1.0 pg TEQ/kg BW/d (Germany) and 3.5 pg TEQ/kg BW/d (Spain) [31-34]. In the US, dietary dioxin exposure was estimated between 0.3 and 3.2 pg TEQ/kg BW/d [35]. Dairy products and seafood were consistently identified across studies as being the main food contributors to dioxin exposure [21,32,34]. However, because food contribution depends on dietary habits, some studies (including ours) showed an additional important contribution of cereal products and fruit and vegetables [31]. These foods, although little contaminated by dioxins, are highly consumed by some populations, and their contribution to overall dioxin intake must thus be considered. However, most previous studies focused on dioxin intake from foods known to be the most contaminated by dioxins and did not consider dioxin intake from consumption of foods from vegetal origin. This is likely to result in an underestimation of the

intake of dioxins, as concluded in a validation study of a dietary questionnaire developed for the assessment of dietary dioxin exposure [36].

Previous studies investigating the association between breast cancer and dioxin exposure (environmental, accidental and occupational) mostly reported a weak increase in the risk of breast cancer [9,14] or the mortality rates of breast cancer [11-13]. Our results are not consistent with those studies, probably because they have focused on populations exposed accidentally, occupationally or living near emitting sources of dioxins, that is, exposed to high levels of dioxins, whereas our study population was exposed to low levels of dioxins. Moreover, among the Seveso Women's Health Study (SWHS) [9], the results were obtained for a different exposure window than women from the E3N study, as the SWHS women were aged 0 to 40 at the time of exposure (1976) whereas in 1993, almost half of the women of our study population were postmenopausal.

Our results suggest that exposure to dioxins in the low range observed in our study may be associated with a decreased risk of hormone-independent (ER-/PR-) breast cancer, especially after menopause. Inverse associations between endocrine disrupting chemicals and ER- breast cancer risk have been found previously. Raaschou-Nielsen *et al.* found a lower risk associated with higher adipose tissue concentrations of dioxin-like organochlorines in postmenopausal women [37]. Gammon *et al.*

Table 4 Hazard ratios for pre- and postmenopausal breast cancer risk for increased intake of 0.43 pg/kg body weight/day, according to quartiles of estimated dietary dioxin exposure (n = 63,830), 1993 to 2008

Populations	/0.43	Dietary dioxin exposure (pg/kg body weight/day)				P-trend
		<0.98 (ref)	[0.98, 1.23]	[1.23, 1.52]	≥1.52	
Premenopausal women^c						
Cases, n	446	104	105	117	120	
HR (95% CI) ^a	1.03 (0.94, 1.13)	1.00	0.93 (0.71, 1.22)	0.96 (0.74, 1.25)	0.94 (0.73, 1.23)	0.5705
HR (95% CI) ^b	0.98 (0.87, 1.12)	1.00	0.84 (0.63, 1.11)	0.83 (0.62, 1.11)	0.77 (0.54, 1.08)	0.7873
Postmenopausal women^d						
Cases, n	3019	775	749	731	764	
HR (95% CI) ^a	1.01 (0.98, 1.05)	1.00	0.97 (0.88, 1.08)	0.96 (0.87, 1.06)	1.03 (0.93, 1.14)	0.5134
HR (95% CI) ^b	0.98 (0.93, 1.03)	1.00	0.94 (0.84, 1.04)	0.90 (0.80, 1.01)	0.93 (0.82, 1.06)	0.3827
ER+/PR+^e						
Cases, n	1365	354	336	325	350	
HR (95% CI) ^a	1.02 (0.96, 1.07)	1.00	0.96 (0.82, 1.11)	0.94 (0.81, 1.09)	1.03 (0.89, 1.20)	0.5837
HR (95% CI) ^b	0.97 (0.90, 1.04)	1.00	0.91 (0.78, 1.06)	0.86 (0.73, 1.02)	0.91 (0.75, 1.10)	0.4193
ER+/PR-^e						
N cases	513	117	136	125	135	
HR (95% CI) ^a	1.05 (0.96, 1.14)	1.00	1.17 (0.92, 1.50)	1.09 (0.85, 1.40)	1.20 (0.93, 1.53)	0.2780
HR (95% CI) ^b	1.05 (0.93, 1.18)	1.00	1.16 (0.89, 1.49)	1.07 (0.81, 1.41)	1.19 (0.86, 1.63)	0.4341
ER-/PR-^e						
N cases	359	98	91	89	81	
HR (95% CI) ^a	0.95 (0.85, 1.06)	1.00	0.93 (0.70, 1.23)	0.91 (0.68, 1.21)	0.84 (0.62, 1.13)	0.3261
HR (95% CI) ^b	0.86 (0.74, 1.00)	1.00	0.84 (0.63, 1.13)	0.77 (0.56, 1.07)	0.65 (0.45, 0.96)	0.0463

^aAge-adjusted models; ^badjusted for age, height, body mass index, energy intake, education, physical activity, smoking status, menopausal status combined with use of menopausal hormone treatment, alcohol intake, age at menarche, use of oral contraceptives, use of progestin, age at menopause, age at first full-term pregnancy and number of live births, breastfeeding, family history of breast cancer, history of personal benign breast disease and mammography; ^cresults for premenopausal breast cancer risk according to hormone receptor status were not presented due to the small number of cases; ^dER-/PR+ postmenopausal invasive breast cancer risk was not analyzed due to small number of cases (n = 73); ^en = 709 postmenopausal invasive breast cancer cases with missing hormone receptor status. HR, hazard ratio.

observed a lower risk of ER-/PR- breast cancer in women with serum levels of PCBs in the highest tertile compared to the lowest [38]. Some studies investigating breast cancer risk in relation to dioxin exposure also suggested a decreased breast cancer risk. A French registry-based case-control study conducted near a MSWI showed a decreased risk of invasive breast cancer among women aged over 60 years and highly exposed, compared to non-exposed women, although these results could be explained by confounding, because there was no adjustment on breast cancer risk factors or other potential confounders other than age and place of residence [15]. A small decrease in breast cancer risk was also observed among women exposed during the Seveso accident and followed between 1977 and 1986, although not statistically significant [7]. However, the authors of these studies did not stratify on tumor hormone receptor status. Our findings may also be supported by experimental evidence on the role of dioxins in promotion or inhibition of mammary tumor formation. TCDD has been shown to operate in animals and humans through

binding to the aryl hydrocarbon receptor (AhR, [5]). Some studies have described an antiproliferative action of TCDD in ER- breast cancer cell lines through AhR-dependent [39,40] or AhR-independent pathways [41]. These observations were complemented by animal studies demonstrating that TCDD inhibits the proliferation of cells from the mammary gland and mammary tumor formation [42,43]. In a recent review of the literature on the role of AhR in carcinogenesis, Safe *et al.* highlighted the fact that the effect of TCDD may vary according to the target organ as well as windows of exposure [44]. For example, animal studies have demonstrated that prenatal TCDD exposure increased susceptibility to carcinogen-induced mammary tumor formation, while exposure during pregnancy delayed mammary tumor formation. Women of our study population were on average aged 57.2 ± 6.6 years old at inclusion in 1993. Therefore, we were not able to investigate the association of breast cancer risk with estimated dietary dioxin exposure during other windows of exposure such as prenatal periods or pregnancy.

Strengths and limitations

Strengths of our study include the prospective design, the large size of the study population and number of breast cancer cases, a long follow-up that provides a large latency period for potential cancer occurrence, adjustment for most known breast cancer risk factors updated throughout follow-up, and validated dietary data. The participation rate which is still high 18 years after the start of the E3N study ($\geq 75\%$ at each follow-up questionnaire) and the low percentage of missing data attests of the quality of the information collected. Because our statistical models were adjusted for several individual breast cancer risk factors, it is unlikely that residual confounding changes our risk estimates. Moreover, we analyzed the potential confounding effect of flavonoid intake. Flavonoids are dietary phytochemicals that have been suggested to protect against dioxin toxicity [45] and to play a role in breast cancer prevention [46,47]; adjustment for total dietary flavonoid intake did not change our results.

Nevertheless, our study has some limitations. First, to determine dietary dioxin exposure, we used consumption data obtained through a semi-quantitative diet history questionnaire. DHQ are often used for their applicability to a large number of participants. However, as DHQ are self-administered questionnaires on habitual food consumption over the past 12 months, we cannot exclude some inaccuracy of consumption data and a non-differential bias due to the desire of conformity to social norms. Nevertheless, the potential lack of precision has been minimized by the collection of food frequency for each food item and the estimation of portion sizes through a photo-booklet. Moreover, both the dietary questionnaire and photo-booklet have been validated [18,19].

Second, the food contamination data by dioxins came from the first French national monitoring program, which aimed to assess dioxin levels of foods sampled between 1996 and 1998 [21]. Because the authors did not achieve an exhaustive sampling plan, the data may not be representative of the dioxin contamination of food in France for that period. However, because of the large number of samples ($n = 444$), the coverage of all major food contributors of dioxin intake and the fact that those data were the closest to the period of the E3N DHQ administration, we can assume that they accurately reflect the contamination of food in 1993. We combined the consumption data with the contamination data using standardized formulas recommended by the WHO [22]. However, for some food items of the E3N DHQ, no contamination data were reported in the CSHPF study; therefore we assigned to each of these items the average contamination of the food group to which it belonged (see Additional file 1: Table S1 for an example).

Also, we could not consider the source of the food consumed, that is, women living and consuming food

produced near a dioxin emitting source may consume foods more contaminated than average. This could have led to an underestimation of the exposure for those women and to misclassification bias, although non differential so the risk would be potentially drawn toward 1. The estimated dietary dioxin exposure may not reflect the total exposure to dioxins of the E3N women. We cannot eliminate the contribution of exposure from environmental sources, even if its contribution remains probably minor compared to that of food. In contrast, because most women from the E3N cohort are teachers or from affiliated occupations, we can assume the occupational dioxin exposure to be negligible and homogeneous among study participants.

Our estimation of the dietary dioxin exposure cannot be extrapolated to the French general population. Indeed, the intake of dioxins through the diet depends on the relative intake of foods of high or low levels of dioxin contamination and their quantity consumed (for instance, our population has a higher consumption of fruit and vegetables than generally observed). Moreover, we can expect a healthy participation effect in our cohort, as it is composed of volunteers with a high level of education and health consciousness. This could result in a possible underrepresentation of high fat (so high dioxin) consumptions; however this should not have biased our results.

Conclusion

Among women from the E3N prospective cohort, estimated dietary dioxin exposure was not associated with breast cancer risk. Further studies should include both dietary and non-dietary (environmental) exposures as well as different windows of exposure (prenatal, pregnancy periods) in order to further investigate the complex association between low-dose dioxin exposure and hormone receptor-defined breast cancer risk.

Additional file

Additional file 1: Table S1. Dioxin contamination and lipids proportion of food items from the E3N diet history questionnaire of 1993. **Table S2.** Hazards ratios (HR) for invasive breast cancer in subgroups defined by body mass index, weight change, dietary patterns and breastfeeding ($n = 63,830$), 1993 to 2008.

Abbreviations

AhR: aryl hydrocarbon receptor; BMI: body mass index; BW: body weight; CSHPF: *Conseil Supérieur d'Hygiène Publique de France*; DHQ: diet history questionnaire; E3N: *Étude Épidémiologique auprès de femmes de la Mutuelle Générale de l'Éducation Nationale*; ER: estrogen receptor; HR: hazard ratio; MSWI: municipal solid-waste incinerator; PCDD: polychlorinated dibenzo-*p*-dioxin; PCDF: polychlorinated dibenzofuran; PR: Progesterone receptor; TCDD: 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; TEQ: toxic equivalency; WHO: World Health Organization.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

FCC and MCBR have contributed to the conception and design of the E3N study. BF, TP and LD conceived the study on dioxin exposure and breast cancer risk and participated in its design. FCC, MCBR and LD have made contribution to the data acquisition. AMND performed the statistical analyses. AMND drafted the manuscript. AMND, LD and BF have interpreted the results. LD, BF, FCC, MCBR and TP have made contribution to the revision of the manuscript. All authors have read and approved the submitted manuscript.

Acknowledgements

We gratefully acknowledge the participants for providing data and the practitioners for providing pathology reports. We are thankful to Jean-Luc Volatier for his advice on dioxin exposure calculation. We thank Marina Touillaud for her assistance on obtaining data on flavonoids intake and Raul Zamora-Ros for providing these data. The E3N study is financially supported by the French League Against Cancer, the *Mutuelle Générale de l'Education Nationale*, the *Institut Gustave Roussy*, the *Institut National de la Santé et de la Recherche Médicale*. The present project also received grant from the Oncostarter Program of the *Cancéropôle Lyon Auvergne Rhône-Alpes* and the French Environment and Energy Management Agency. Aurélie Danjou is supported by a doctoral fellowship from the laboratory *Santé, Individu, Société, Ecole Doctorale Interdisciplinaire Sciences-Santé, Université Claude Bernard Lyon 1*.

Author details

¹Unité Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon, Cedex 08, France. ²Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69100 Villeurbanne, France. ³Inserm, Center for Research in Epidemiology and Population Health (CESP), U1018, Nutrition, Hormones and Women's Health team, 114 rue Edouard-Vaillant, 94805 Villejuif Cedex, France. ⁴Université Paris-Sud, UMRS 1018, Villejuif, France. ⁵Gustave Roussy, F-94805 Villejuif, France.

Received: 10 September 2014 Accepted: 11 February 2015

Published online: 17 March 2015

References

- Fries GF. A review of the significance of animal food products as potential pathways of human exposures to dioxins. *J Anim Sci*. 1995;73:1639–50.
- Kogevinas M. Human health effects of dioxins: cancer, reproductive and endocrine system effects. *Hum Reprod Update*. 2001;7:331–9.
- Liem AK, Furst P, Rappe C. Exposure of populations to dioxins and related compounds. *Food Addit Contam*. 2000;17:241–59.
- Schug TT, Janesick A, Blumberg B, Heindel JJ. Endocrine disrupting chemicals and disease susceptibility. *J Steroid Biochem Mol Biol*. 2011;127:204–15.
- Baan R, Grosse Y, Straif K, Secretan B, El GF, Bouvard V, et al. A review of human carcinogens—Part F: chemical agents and related occupations. *Lancet Oncol*. 2009;10:1143–4.
- U.S. EPA. Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-P-dioxin (TCDD) and Related Compounds National Academy Sciences (External Review Draft). Washington, D.C.: U.S. Environmental Protection Agency; 2004. EPA/600P-00/001Cb.
- Bertazzi A, Pesatori AC, Consonni D, Tironi A, Landi MT, Zocchetti C. Cancer incidence in a population accidentally exposed to 2,3,7,8-tetrachlorodibenzo-para-dioxin. *Epidemiology*. 1993;4:398–406.
- Pesatori AC, Consonni D, Rubagotti M, Grillo P, Bertazzi PA. Cancer incidence in the population exposed to dioxin after the "Seveso accident": twenty years of follow-up. *Environ Health*. 2009;8:39.
- Warner M, Eskenazi B, Mocarelli P, Gerthoux PM, Samuels S, Needham L, et al. Serum dioxin concentrations and breast cancer risk in the Seveso Women's Health Study. *Environ Health Perspect*. 2002;110:625–8.
- Warner M, Mocarelli P, Samuels S, Needham L, Brambilla P, Eskenazi B. Dioxin exposure and cancer risk in the Seveso Women's Health Study. *Environ Health Perspect*. 2011;119:1700–5.
- Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. Environmental pollutants and breast cancer: epidemiologic studies. *Cancer*. 2007;109:2667–711.
- Manuwald U, Velasco GM, Berger J, Manz A, Baur X. Mortality study of chemical workers exposed to dioxins: follow-up 23 years after chemical plant closure. *Occup Environ Med*. 2012;69:636–42.
- Revich B, Aksel E, Ushakova T, Ivanova I, Zhuchenko N, Klyuev N, et al. Dioxin exposure and public health in Chapaevsk. *Russia Chemosphere*. 2001;43:951–66.
- Fabre P, Daniau C, Gorla S, de Crouy-Chanel P, Empereur-Bissonnet P. Etude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères [in French]. In: Rapport d'étude et synthèse. Saint-Maurice: Institut de veille sanitaire; 2008. p. 136.
- Viel JF, Clement MC, Hagi M, Grandjean S, Challier B, Danzon A. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure. *Int J Health Geogr*. 2008;7:4.
- Clavel-Chapelon F, van Liere MJ, Giubout C, Niravong MY, Goulard H, Le CC, et al. E3N, a French cohort study on cancer risk factors. E3N Group. *Etude Epidémiologique auprès de femmes de l'Education Nationale*. *Eur J Cancer Prev*. 1997;6:473–8.
- CépiDc. Centre d'épidémiologie sur les causes médicales de décès. <http://www.cepidc.inserm.fr/site4/>.
- van Liere MJ, Lucas F, Clavel F, Slimani N, Villemainot S. Relative validity and reproducibility of a French dietary history questionnaire. *Int J Epidemiol*. 1997;26:5128–36.
- Lucas F, Niravong M, Villemainot S, Kaaks R, Clavel-Chapelon F. Estimation of food portion size using photographs: validity, strengths, weaknesses and recommendations. *J Hum Nutr Diet*. 1995;8:65–74.
- Favier J, Ireland-Ripert J, Toque C, Feinberg M. Répertoire général des aliments: Table de composition. Composition tables. CIQUAL-REGAL. Paris, France: INRA Editions; 1995. French.
- CSHPF. Dioxines: données de contamination et d'exposition de la population française. France: Agence française de sécurité sanitaire des aliments (Afssa); 2000. French.
- FAO/WHO. Dietary exposure assessment of chemicals in food. Report of a joint FAO/WHO consultation. World Health Organization and the Food Agriculture Organization of the United Nations: Annapolis, MA; 2005.
- NATO. Scientific basis for the development of the international toxicity equivalency factor (i-tef) method of risk assessment for complex mixtures of dioxins and related compounds. committee on the challenges of modern society. 178. Washington, D.C.: U.S. Environmental Protection Agency; 1988.
- JEFCFA. Joint FAO/WHO Expert Committee on Food Additives, fifty-seventh meeting. summary and conclusions. World Health Organization and the Food Agriculture Organization of the United Nations: Rome; 2001.
- van Leeuwen FX, Feeley M, Schrenk D, Larsen JC, Farland W, Younes M. Dioxins: WHO's tolerable daily intake (TDI) revisited. *Chemosphere*. 2000;40:1095–101.
- Tehard B, Lahmann PH, Riboli E, Clavel-Chapelon F. Anthropometry, breast cancer and menopausal status: use of repeated measurements over 10 years of follow-up-results of the French E3N women's cohort study. *Int J Cancer*. 2004;111:264–9.
- Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F, et al. Postmenopausal breast cancer risk and dietary patterns in the E3N-EPIC prospective cohort study. *Am J Epidemiol*. 2009;170:1257–67.
- Schechter A, Ryan JJ, Papke O. Decrease in levels and body burden of dioxins, dibenzofurans, PCBS, DDE, and HCB in blood and milk in a mother nursing twins over a thirty-eight month period. *Chemosphere*. 1998;37:1807–16.
- Uemura H, Arisawa K, Hiyoshi M, Satoh H, Sumiyoshi Y, Morinaga K, et al. PCDDs/PCDFs and dioxin-like PCBs: recent body burden levels and their determinants among general inhabitants in Japan. *Chemosphere*. 2008;73:30–7.
- Arisawa K, Uemura H, Hiyoshi M, Kitayama A, Takami H, Sawachika F, et al. Dietary patterns and blood levels of PCDDs, PCDFs, and dioxin-like PCBs in 1656 Japanese individuals. *Chemosphere*. 2011;82:656–62.
- Domingo JL, Schuhmacher M, Granero S, Llobet JM. PCDDs and PCDFs in food samples from Catalonia, Spain. An assessment of dietary intake. *Chemosphere*. 1999;38:3517–28.
- Harrison N, Wearne S, Gem MG, Gleadle A, Startin J, Thorpe S, et al. Time trends in human dietary exposure to PCDDs, PCDFs and PCBs in the UK. *Chemosphere*. 1998;37:1657–70.
- Jimenez B, Hernandez LM, Eljarrat E, Rivera J, Gonzalez MJ. Estimated intake of PCDDs, PCDFs and co-planar PCBs in individuals from Madrid (Spain) eating an average diet. *Chemosphere*. 1996;33:1465–74.
- Malisch R. Update of PCDD/PCDF-intake from food in Germany. *Chemosphere*. 1998;37:1687–98.

35. Schecter A, Startin J, Wright C, Kelly M, Papke O, Lis A, et al. Dioxins in U.S. food and estimated daily intake. *Chemosphere*. 1994;29:2261–5.
36. Bilau M, Matthys C, Bellemans M, De NM, Willems JL, De HS. Reproducibility and relative validity of a semi-quantitative food frequency questionnaire designed for assessing the intake of dioxin-like contaminants. *Environ Res*. 2008;108:327–33.
37. Raaschou-Nielsen O, Pavuk M, Leblanc A, Dumas P, Philippe WJ, Olsen A, et al. Adipose organochlorine concentrations and risk of breast cancer among postmenopausal Danish women. *Cancer Epidemiol Biomarkers Prev*. 2005;14:67–74.
38. Gammon MD, Wolff MS, Neugut AJ, Eng SM, Teitelbaum SL, Britton JA, et al. Environmental toxins and breast cancer on Long Island. II. Organochlorine compound levels in blood. *Cancer Epidemiol Biomarkers Prev*. 2002;11:686–97.
39. Wang WL, Porter W, Burghardt R, Safe SH. Mechanism of inhibition of MDA-MB-468 breast cancer cell growth by 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Carcinogenesis*. 1997;18:925–33.
40. Zhang S, Lei P, Liu X, Li X, Walker K, Kotha L, et al. The aryl hydrocarbon receptor as a target for estrogen receptor-negative breast cancer chemotherapy. *Endocr Relat Cancer*. 2009;16:835–44.
41. Yoshioka H, Hiromori Y, Aoki A, Kimura T, Fujii-Kuriyama Y, Nagase H, et al. Possible aryl hydrocarbon receptor-independent pathway of 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced antiproliferative response in human breast cancer cells. *Toxicol Lett*. 2012;211:257–65.
42. Lew BJ, Collins LL, O'Reilly MA, Lawrence BP. Activation of the aryl hydrocarbon receptor during different critical windows in pregnancy alters mammary epithelial cell proliferation and differentiation. *Toxicol Sci*. 2009;111:151–62.
43. Kociba RJ, Keyes DG, Beyer JE, Carreon RM, Wade CE, Dittenber DA, et al. Results of a two-year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in rats. *Toxicol Appl Pharmacol*. 1978;46:279–303.
44. Safe S, Lee SO, Jin UH. Role of the aryl hydrocarbon receptor in carcinogenesis and potential as a drug target. *Toxicol Sci*. 2013;135:1–16.
45. Ashida H, Fukuda I, Yamashita T, Kanazawa K. Flavones and flavonols at dietary levels inhibit a transformation of aryl hydrocarbon receptor induced by dioxin. *FEBS Lett*. 2000;476:213–7.
46. Kandaswami C, Lee LT, Lee PP, Hwang JJ, Ke FC, Huang YT, et al. The antitumor activities of flavonoids. *In Vivo*. 2005;19:895–909.
47. Moon YJ, Wang X, Morris ME. Dietary flavonoids: effects on xenobiotic and carcinogen metabolism. *Toxicol In Vitro*. 2006;20:187–210.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Additional File 1

Table S1. Dioxin contamination and lipids proportion of food items from the E3N diet history questionnaire of 1993.

Food groups	Food items	Dioxin contamination level (pg I-TEQ/g lipid weight or <i>fresh weight</i>)^a	Lipids proportion (%)^c
Meat	Beef	0.80	8.86
	Horse	0.60 ^b	4.60
	Lamb	0.74	14.09
	Calf	0.48	5.12
	Pork	0.16	11.04
	Offal	3.32	15.42
	Charcuterie	0.25	16.41
	Poultry	0.60	6.76
	Rabbit	0.60 ^b	8.70
Cereal products	Pasta	0.01	-
	Rice	0.01	-
	Cereals	0.02	-
	Bread	0.01	-
	Biscuits/cakes	0.0125	-
Eggs	Eggs	1.51	10.50
Fish and Seafood	Trout	9.36	4.20
	Sardines	2.70	12.30
	Colin/hake	2.24	1.00
	Ling	4.64 ^b	1.00
	Mackerel	3.23	20.90
	Flounder	4.64 ^b	1.00
	Haddock	4.64 ^b	1.00
	Sole	4.64 ^b	1.00
	Saithe	3.38	1.00
	Whiting	34.16	0.90
	Cod	8.45	0.95
	Salmon	5.15	10.78
	Saumonette	4.64 ^b	16.30
	Other fish	5.13	4.18
	Canned fish	4.64 ^b	8.50
	Seafood	30.22	2.77
Dairy products	Milk	0.65	2.32
	Butter	0.92	68.67
	Pressed cheese	0.64	29.25
	Other cheese	0.77	30.10
	Cream	0.68	27.00
	Fromage blanc	1.16	3.35

	Dairy dessert	1.16	5.04
	Yogurt	1.16	1.80
Fruit and vegetables	Fruits	0.01	-
	Leafy vegetables	0.055	-
	Root vegetables	0.01	-
	Other vegetables	0.032	-
	Soup	0.034 ^b	-
Added fats	Animal fats	0.88	99.75
	Peanut oil	0.03	83.25
	Olive oil	0.04	83.25
	Margarine	0.04	61.99
	Sunflower oil	0.04 ^b	83.25
	Corn oil	0.04 ^b	83.25
	Other vegetable fats	0.04 ^b	70.75
	Mayonnaise	0.04 ^b	78.60

^a Dioxin contamination data of 1996-1998 provided by the French High Council for Public

5 Health (Conseil Supérieur d'Hygiène Publique de France).

^b The average dioxin contamination of the food group was assigned to the food item.

^c Lipids proportion provided by the food composition table developed by the E3N team.

Table S2. Hazards ratios (HR) for invasive breast cancer in subgroups defined by body mass

10 index, weight change, dietary patterns and breastfeeding (N=63,830), 1993-2008.

Stratifications	Dietary dioxin exposure (pg/kg body weight/day)					p-trend
	/0.43	<0.98 (ref)	[0.98,1.23[[1.23,1.52[≥1.52	
Body Mass Index (BMI)						
BMI<25 kg/m²						
N cases	2778	600	684	706	788	
HR (95% CI) ^a	1.01 (0.97, 1.05)	1.00	0.99 (0.89, 1.11)	0.96 (0.86, 1.07)	1.01 (0.91, 1.13)	0.5931
HR (95% CI) ^b	0.98 (0.93, 1.03)	1.00	0.96 (0.86, 1.07)	0.90 (0.80, 1.02)	0.93 (0.81, 1.06)	0.4682
BMI≥25 kg/m²						
N cases	687	280	169	142	96	
HR (95% CI) ^a	1.01 (0.93, 1.10)	1.00	0.87 (0.72, 1.06)	0.99 (0.81, 1.21)	0.99 (0.78, 1.25)	0.7633
HR (95% CI) ^b	1.00 (0.89, 1.12)	1.00	0.85 (0.69, 1.04)	0.94 (0.75, 1.18)	0.94 (0.70, 1.26)	0.9875
Weight change^c						
Weight loss						
(>-2 kg/5 years)						
N cases	211	56	59	47	49	
HR (95% CI) ^a	1.03 (0.91, 1.18)	1.00	1.19 (0.83, 1.72)	1.06 (0.72, 1.57)	1.12 (0.77, 1.65)	0.6288
HR (95% CI) ^b	1.00 (0.84, 1.19)	1.00	1.11 (0.76, 1.63)	1.00 (0.65, 1.54)	1.01 (0.62, 1.64)	0.9796
Stable weight						
([-2;2] kg/5 years)						
N cases	2014	501	482	498	533	
HR (95% CI) ^a	1.00 (0.95, 1.04)	1.00	0.90 (0.80, 1.02)	0.89 (0.79, 1.01)	0.95 (0.84, 1.07)	0.9628
HR (95% CI) ^b	1.00 (0.94, 1.07)	1.00	0.89 (0.78, 1.01)	0.88 (0.76, 1.01)	0.92 (0.79, 1.08)	0.9213
Weight gain						
(>2 kg/5 years)						
N cases	1234	320	310	303	301	
HR (95% CI) ^a	1.04 (0.99, 1.10)	1.00	1.04 (0.89, 1.22)	1.09 (0.93, 1.27)	1.13 (0.96, 1.32)	0.1350
HR (95% CI) ^b	1.00 (0.93, 1.08)	1.00	1.00 (0.85, 1.18)	1.00 (0.84, 1.19)	1.02 (0.83, 1.25)	0.9893
Western Pattern						
Low intake						
N cases	1059	483	299	198	79	
HR (95% CI) ^a	0.99 (0.91, 1.07)	1.00	1.03 (0.89, 1.19)	1.12 (0.95, 1.33)	0.96 (0.75, 1.22)	0.7361
HR (95% CI) ^b	0.98 (0.88, 1.08)	1.00	1.03 (0.89, 1.20)	1.12 (0.94, 1.34)	0.97 (0.76, 1.24)	0.6586
Medium intake						

N cases	1163	274	321	349	219	
HR (95% CI) ^a	0.96 (0.89, 1.03)	1.00	0.87 (0.74, 1.02)	0.95 (0.81, 1.11)	0.81 (0.67, 0.96)	0.2150
HR (95% CI) ^b	0.94 (0.86, 1.02)	1.00	0.85 (0.73, 1.01)	0.93 (0.79, 1.09)	0.78 (0.65, 0.95)	0.1438
High intake						
N cases	1243	110	222	386	525	
HR (95% CI) ^a	0.98 (0.93, 1.03)	1.00	0.92 (0.73, 1.15)	0.94 (0.76, 1.16)	0.85 (0.69, 1.05)	0.4446
HR (95% CI) ^b	1.02 (0.96, 1.09)	1.00	0.94 (0.74, 1.18)	0.98 (0.78, 1.21)	0.89 (0.72, 1.11)	0.4666
Healthy pattern						
Low intake						
N cases	1166	375	292	287	212	
HR (95% CI) ^a	1.01 (0.94, 1.08)	1.00	0.96 (0.82, 1.12)	1.06 (0.91, 1.24)	0.88 (0.74, 1.05)	0.8799
HR (95% CI) ^b	1.00 (0.91, 1.10)	1.00	0.93 (0.80, 1.09)	1.02 (0.86, 1.20)	0.83 (0.69, 1.01)	0.9751
Medium intake						
N cases	1164	290	292	319	263	
HR (95% CI) ^a	0.98 (0.92, 1.05)	1.00	1.00 (0.85, 1.18)	1.05 (0.90, 1.23)	0.92 (0.78, 1.10)	0.6134
HR (95% CI) ^b	1.00 (0.91, 1.09)	1.00	1.00 (0.85, 1.18)	1.04 (0.87, 1.23)	0.90 (0.74, 1.09)	0.9096
High intake						
N cases	1135	202	258	327	348	
HR (95% CI) ^a	1.06 (1.01, 1.12)	1.00	1.08 (0.90, 1.30)	1.24 (1.04, 1.48)	1.23 (1.03, 1.46)	0.0248
HR (95% CI) ^b	1.05 (0.98, 1.13)	1.00	1.06 (0.88, 1.28)	1.21 (1.00, 1.46)	1.17 (0.95, 1.43)	0.1574
Breastfeeding						
No breastfeeding						
N cases	1483	382	366	371	364	
HR (95% CI) ^a	1.02 (0.97, 1.07)	1.00	1.00 (0.87, 1.15)	1.03 (0.89, 1.19)	1.02 (0.88, 1.18)	0.4492
HR (95% CI) ^b	1.02 (0.95, 1.09)	1.00	0.97 (0.84, 1.13)	0.99 (0.84, 1.16)	0.99 (0.82, 1.19)	0.6577
Breastfeeding						
N cases	1982	497	488	477	520	
HR (95% CI) ^a	1.01 (0.97, 1.06)	1.00	0.95 (0.84, 1.07)	0.92 (0.81, 1.04)	1.01 (0.89, 1.14)	0.6589
HR (95% CI) ^b	0.99 (0.93, 1.05)	1.00	0.92 (0.81, 1.05)	0.88 (0.76, 1.01)	0.94 (0.80, 1.11)	0.7427
Smoking Status						
Non-smoker						
N cases	1861	492	469	431	469	
HR (95% CI) ^a	1.00 (0.96, 1.05)	1.00	0.96 (0.84, 1.09)	0.88 (0.77, 1.00)	0.99 (0.87, 1.12)	0.9040
HR (95% CI) ^b	0.99 (0.92, 1.05)	1.00	0.93 (0.81, 1.06)	0.83 (0.72, 0.96)	0.92 (0.78, 1.09)	0.6414
Former smoker						
N cases	1158	278	273	309	298	
HR (95% CI) ^a	1.03 (0.97, 1.09)	1.00	0.95 (0.81, 1.13)	1.09 (0.92, 1.28)	1.05 (0.89, 1.23)	0.2918

	HR (95% CI) ^b	1.03 (0.95, 1.11)	1.00	0.94 (0.79, 1.12)	1.07 (0.89, 1.28)	1.03 (0.84, 1.27)	0.4625
Smoker							
	N cases	446	109	112	108	117	
	HR (95% CI) ^a	1.01 (0.93, 1.11)	1.00	1.07 (0.82, 1.39)	1.03 (0.79, 1.34)	1.05 (0.81, 1.36)	0.7549
	HR (95% CI) ^b	1.00 (0.88, 1.13)	1.00	1.03 (0.78, 1.36)	0.97 (0.72, 1.30)	0.98 (0.70, 1.37)	0.9852
Alcohol intake							
0 g/day							
	N cases	375	134	90	68	83	
	HR (95% CI) ^a	1.08 (0.98, 1.19)	1.00	0.86 (0.66, 1.13)	0.78 (0.58, 1.05)	1.21 (0.92, 1.59)	0.1448
	HR (95% CI) ^b	1.02 (0.89, 1.17)	1.00	0.79 (0.60, 1.04)	0.68 (0.49, 0.94)	0.99 (0.69, 1.42)	0.7708
<6.9 g/day							
	N cases	1234	348	338	294	254	
	HR (95% CI) ^a	0.99 (0.93, 1.05)	1.00	1.02 (0.88, 1.19)	0.96 (0.82, 1.12)	0.96 (0.81, 1.12)	0.6697
	HR (95% CI) ^b	1.00 (0.92, 1.08)	1.00	1.03 (0.88, 1.20)	0.96 (0.81, 1.14)	0.96 (0.78, 1.19)	0.9285
≥6.9 g/day							
	N cases	1856	397	426	486	547	
	HR (95% CI) ^a	1.00 (0.96, 1.05)	1.00	0.94 (0.82, 1.07)	0.96 (0.84, 1.10)	0.97 (0.86, 1.11)	1.0000
	HR (95% CI) ^b	1.01 (0.95, 1.07)	1.00	0.92 (0.80, 1.06)	0.95 (0.82, 1.11)	0.97 (0.82, 1.15)	0.8817

^a age-adjusted models.

^b adjusted for age, height, body mass index, energy intake, education, physical activity, smoking status combined with use of menopausal hormone treatment, alcohol intake, age at menarche, use of oral contraceptives, use of progestin, age at menopause, age at first full-term pregnancy and number of live births, breastfeeding, family history of breast cancer, history of personal benign breast disease and mammography.

^c n=6 breast cancer cases with missing data for weight change during follow-up.

3. Conclusion

This was the first study investigating the association between dietary dioxin exposure and breast cancer risk, and overall no increase in risk was observed among women from the E3N cohort. Factors such as dietary patterns, weight change and breastfeeding did not influence the association. Further adjustment for flavonoid intake did not change the results. However, menopausal status and hormone receptor status may modify the association: a significant decreased risk of ER-negative/PR-negative breast cancer was found among post-menopausal women, when comparing women with low *versus* high dioxin intake.

Moreover, the latter observation was consistent with previous studies reporting inverse associations between EDCs and ER-negative breast cancer risk (Gammon et al. 2002b; Raaschou-Nielsen et al. 2005), and a decreased risk of breast cancer was reported among French women over 60 years of age and highly exposed to dioxins in the vicinity of a MSWI (Viel et al. 2008). This raised a different hypothesis about the dioxin mechanism at low-doses that could lead to the decrease of hormone-independent breast cancer risk. Experimental studies also supported these results, highlighting the anti-proliferative action of TCDD on *in vitro* breast cancer cells (Wang et al. 1997; Zhang et al. 2009), animal mammary gland (Kociba et al. 1978; Lew et al. 2009) and rodents (M Hall 2013; Safe et al. 2013). Furthermore, other AhR agonists have recently been identified to prevent growth and promote regression of ER-negative breast cancers (Safe et al. 2013).

Major limitations of the study were related to food contamination data sampling and the origin of food and animal consumed. Because the food contamination data came from the first French national program (CSHPF 2000), they were not collected with an exhaustive sampling plan and did not cover every food item of the DHQ3. Monitoring programs that followed, in 2001-2004 and 2006-2010, were based on Total Diet Studies (TDS), a standardized method recommended by WHO to assess dietary exposure to food contaminants. Data from the first French TDS may be more complete and accurate due to an improved sampling plan (Tard et al. 2007); however, they cannot represent the dioxin contamination of 1993, mostly because of the strong decrease in dioxin emission observed in France since 1990, leading to a decrease in dioxin contamination of food. It is though possible to re-assess dietary dioxin exposure in the E3N cohort, using the food consumption data from the E3N DHQ5 of 2005 and the food contamination data from the TDS; investigation of the risk of breast cancer associated with the updated dietary dioxin exposure will be possible in future research with a longer follow-up of the E3N cohort.

The inability to consider in our study the origin of food and animal consumed could have led to an underestimation of the dietary exposure among women living near dioxin emitting sources who could be exposed to local foods more contaminated than average, and to a misclassification bias, although non differential so the risk estimate would be drawn toward unity. Assessing dietary dioxin exposure considering the aspect of food production and farming location could improve accuracy of the exposure and reduce potential misclassification bias.

The study results confirmed the need to consider hormone receptor status of breast tumors and specific windows of breast susceptibility (here, menopause) in epidemiological studies investigating dioxin exposure. Research is needed to clarify the estimation of dietary dioxin exposure regarding the origin of food consumed. Dietary dioxin exposure is considered the main route of dioxin exposure (Hattemer-Frey and Travis 1989; IARC 2012c); however airborne dioxin exposure should not be neglected, in particular for women living in the vicinity of dioxin emitting sources and because of the tendency of dioxins to accumulate in the body.

CHAPTER III. CHARACTERIZATION OF AIRBORNE DIOXIN EXPOSURE

Development of a GIS-based airborne dioxin exposure metric

1. Introduction

Current evidence on the association between dioxin exposure and breast cancer is inconclusive, and methodological limitations, in particular regarding the exposure assessment, have been suggested to partly explain the observed inconsistencies (Brody et al. 2007; Cordioli et al. 2013). Exposure to environmental pollutants, including dioxins, is defined by its intensity, its spatial distribution and temporal components including both duration and variability of the exposure (Cordioli et al. 2013). An important challenge in environmental epidemiological studies on breast cancer is to precisely characterize environmental exposure considering all these three dimensions while also taking into account breast cancer latency and timing regarding breast cancer windows of susceptibility (Brody et al. 2007).

Environmental exposure can be directly assessed with biological measurements that allow integrating individual exposures from all sources (ingestion, inhalation and dermal absorption). However, chemicals may not be measured in the relevant target (the mammary gland; in adipose tissue) and representative for assessing past exposures over longer time periods. Moreover, for some pollutants, existing technologies remain too expensive to be applied in large study populations. Epidemiological studies can also rely on a variety of indirect methods for the assessment of environmental exposure. They include job history, located ambient air quality monitoring networks and surrogates for exposure, such as traffic density or residential proximity metrics (Brody et al. 2007; Cordioli et al. 2013). More complex methods have also been developed to adequately represent the spatial variation of pollutants, such as GIS, land use regression (LUR) models and atmospheric dispersion modeling (de Hoogh et al. 2014).

Major limitations in published studies on the assessment of dioxin exposure are related to the lack of past residential history, historical exposure estimates and individual exposure assessment, as well as the consideration of dioxin exposure from incinerators only and the use of simple surrogates, such as distance from facilities, as proxy for dioxin exposure (Brody et al. 2007; Cordioli et al. 2013). Establishing an association between exposure and disease risk requires assessing exposure at the individual level, which is not possible with an ecological study design and comparison of contaminated areas. Although dioxin emission is expected to decrease with distance from the source, the use of simple residence-to-source distance as a proxy for dioxin concentration could lead to misclassification bias, because dioxin exposure may depend on characteristics of the emission sources (stack height and emission intensity) and on the local environment (meteorology and topography) (Cordioli et al. 2013). Considering participants' residence history is an important aspect as it allows integrating time in the exposure assessment by retracing residential mobility and identifying time lags between exposure and diagnosis (Brody et al. 2007; Han et al. 2013). Long-term dioxin exposure may also vary according to changes in exposure intensity, in particular as a consequence of changes in technologies (Cordioli et al. 2013). Most of published studies that investigated dioxin exposure in relation to breast cancer focused on incinerators, which have been pointed out as one of the most important sources of dioxin emissions. However, other facilities, including metal industries, cement kilns and power plants, are sources likely to emit dioxins and should be considered in the exposure assessment (Nzihou et al. 2012).

Assessing long-term airborne dioxin exposure at the individual address level was the primary challenge of the GEO3N project. Modeling of the airborne dioxin exposure with atmospheric dispersion models was difficult to achieve in our study, due to the nearly 11,110 subjects and 2,000 industrial sources over a 19-year study period, and the large territory to be evaluated (French metropolitan territory). Measurement data were too rare over the study period to allow the use of land use regression models, and the sharp reduction of dioxin emission over the study period prevented us

from considering a recent year as representative of the exposure. On the contrary, GIS methods allowed the development of exposure metrics at fine spatio-temporal scales and over large areas, to estimate retrospective environmental exposure at the individual subject's address with geocoding of residential history. The use of GIS in epidemiological studies to localize precisely residences and dioxin emission sources also allows reducing potential misclassification bias. Moreover, meteorological and topographical parameters can be integrated into a GIS, as well as several types of dioxin emitting sources (Gulliver and Briggs 2011; Han et al. 2013; Hoek et al. 2001; Zou et al. 2009).

The objectives of chapter III were to develop and validate a GIS-based metric to assess airborne dioxin exposure of the women from the GEO3N study population. This work was restricted to the women that lived permanently in the Rhône-Alpes region between 1990 and 2008, with the ultimate goal of applying the metric to the entire GEO3N study population. The Rhône-Alpes region was chosen because of its dense industrial network and the relatively easy access to data.

2. Methods

The airborne dioxin exposure of the GEO3N study participants was assessed at the individual address level, from 1990 to 2008, through a GIS-based method that included (i) the inventory and characterization of the dioxin emitting sources; (ii) the geocoding of participants' residential history and the geolocation of the industrial sources' chimney stack; (iii) the development of a GIS-based metric.

2.1. Inventory and characterization of dioxin emitting sources

2.1.1. Inventory of dioxin emitting sources

Estimation of airborne dioxin exposure of the women from the GEO3N study first required the inventory of industrial sources and quantification of their dioxin releases. A complete retrospective inventory of industrial sources likely to emit or to have emitted dioxins between 1990 and 2008 in France was carried out by two environmental engineers (GH and CD). Industries suspected to be major dioxin emitters were targeted, such as waste incineration, including medical waste incineration, metal production, heat and power generation, production of mineral products, chemicals and consumer goods, and crematoria. The industrial sources were identified through several institutional and public databases, industrial unions and nationally recognized associations and whistleblowers. Structured questionnaires were sent to the facilities identified to collect additional information on their technical characteristics.

2.1.2. Characterization of dioxin emitting sources

Dioxin releases mostly depend on the combustion process and conditions, the type of material burned and the flue gas treatment. Along with inventory, technical characteristics were collected, including geographic location of the chimney stack, operation periods and rates, stack height, process characteristics and flue gas cleaning technologies. In the absence of dioxin emission data from industrial sources over the study period, the Standardized Toolkit for Identification and Quantification of Dioxin and Furan Releases developed by UNEP was applied (UNEP 2013). The Toolkit allowed standardization of dioxin emissions according to activity sectors and operating characteristics and was applicable nationwide. The industrial sources inventoried were thus classified into Toolkit categories and were assigned a dioxin emission factor (in g-TEQ/t). For each distinct operation periods, annual dioxin emission intensity (in g-TEQ/year) was estimated by multiplying the emission factor by the operation rate.

2.2. Geocoding

2.2.1. Geocoding of the participants' residential history

Geocoding is the process of locating and assigning geographic coordinates (X and Y) to subjects' addresses on a map (Han et al. 2013). Along with follow-up questionnaires, the residential history of the E3N women was collected from 1990 to 2008. The address number, address name, zip code and town/municipality were extracted from the questionnaires sent in 1990, 1997, 2000, 2002 and 2005. The consecutive residential addresses (N=28,511 in France including N=4,247 in the Rhône-Alpes region) have been geocoded and included in the GIS in order to locate each woman's place of residence from our study population. The geocoding technician was blinded to the case-control status of the participants.

The use of residential proximity to exposure sources in GIS-based exposure assessment raises concern about the positional accuracy of the subjects' addresses. Studies reported that positional error could lead to exposure misclassification in epidemiological studies (Han et al. 2013). Moreover, while the collection of addresses was achieved prospectively in the E3N cohort, the latter have not been collected for the specific purpose of being geocoded and used in a GIS. Therefore, the choice of the geocoding method should be based in particular on the accuracy of positioning to limit exposure misclassification.

Due to the large number of addresses to be geocoded in the Rhône-Alpes region and on the national territory (N=28,511), we searched for the most accurate automatic method to geocode the residential history of the women from the GEO3N study population. Each consecutive residential address of the women living in the Rhône-Alpes region at recruitment was geocoded *via* three geocoding tools:

- An automatic method based on a free online geocoding service (accessible at www.batchgeocodeur.mapjnz.com/);
- An in-house automatic method using the ArcGIS Software (ArcGIS Locator version 10.0, Environmental System Research Institute – ESRI, Redlands, CA, USA);
- A manual geocoding, considered as the reference method.

The positional accuracy of the residential consecutive addresses was then compared between the two automatic geocoding methods and the manual method. The study was restricted to the women in the Rhône-Alpes region order to apply the best geocoding method to the addresses at the national level. Results and choice of the automatic geocoding method have been described in Article #2.

2.2.2. Geolocation of the dioxin emitting sources

The previous inventory allowed collecting addresses of each industrial source that were then geolocalized in the GIS. This geolocation consisted in a manual positioning at the stack to optimize the accuracy. The positioning at the stack was made through tracking on current and historical satellite images and inquiries from professionals. Different situations have been observed:

- If one stack was found on the industrial site, the point representing the source was positioned at that stack;
- If several stacks were found on one industrial site, the point representing the source was positioned equidistant from each stack;
- If the stack was not visible on satellite images and the industrial site was composed of one building, the point representing the source was located at the centroid of the building;
- If the stack was not visible on satellite images and the industrial site was composed of several buildings, the point representing the source was located at the centroid of the plot;

- If no current and historical images were available, an Internet research was conducted to obtain information allowing the location of the source at the stack or near the industrial site;
- If neither satellite images nor information on the location were available, the point representing the source was located to the town hall of the municipality where the industrial site was implanted.

The accuracy of this manual positioning was recorded for each industrial source when geocoding. The method for geolocating the sources was first applied in the Rhône-Alpes region.

2.3. GIS-based airborne dioxin exposure assessment

Relevant parameters for inclusion in the GIS-based metric were derived from the literature and comparison with dioxin atmospheric dispersion modeling performed in two areas and three distinct periods.

2.3.1. Literature review

A literature review was conducted in order to identify existing methods that developed a metric for the assessment of airborne dioxin exposure, and more generally to air pollution; determine the relevant meteorological and topographical parameters to include in the GIS; and define the GIS-based airborne dioxin metric estimating the GEO3N participants' exposure at the individual address level. PubMed Medline and Scopus bibliographic databases were searched through several algorithms that are presented in Appendice 1. Articles written in English and French were included in the review. The research first focused on epidemiological and methodological studies that used a GIS-based metric to estimate the exposure to dioxins; and was extended to PAHs, dioxin-like compounds, POPs, and globally to air pollution. Secondly, studies that modeled dioxin atmospheric dispersion were identified, along with parameters included in the modeling. Relevant publications were selected through assessment of titles, abstracts and articles by four team members (AD, EF, GH and TC). The literature review was completed with screening of the lists of references from the publications identified previously.

2.3.2. Modeling of airborne dioxin exposure

Prior to the development of the exposure metric, a modeling of the dioxin atmospheric dispersion was performed in restricted areas (by TC), using the SIRANE atmospheric dispersion model. SIRANE is an urban Gaussian dispersion model that integrates a specific module to simulate pollutant dispersion within a built environment that has been developed within the Fluid Mechanics and Acoustics Laboratory (UMR 5509, École Centrale de Lyon) (Soulhac et al. 2012, 2011). SIRANE simulates dispersion and transport of pollutants from one or more emitting sources to one or more receptors (*i.e.* subjects), considering meteorological conditions and geometry of the streets.

An atmospheric dispersion modeling of dioxins was conducted in two areas of the Rhône-Alpes region, Lyon (a non-mountainous urban area) and Le Bugey (a rural area), with distinct dioxin emitting sources and varying emission intensities located in, for three distinct years (1996, 2002 and 2008), with the industrial sources identified through the inventory. Results of the modeling of dioxin atmospheric dispersion in Lyon were comparable to the averaged concentrations provided continuously by a measurement station placed at the city center of Lyon. Then, average dioxin concentrations were calculated for each GEO3N woman from Lyon and Le Bugey at each year, and women were ranked according to their average dioxin concentration and categorized into quintiles. This classification served as reference to determine parameters' combination in the GIS-based metric.

2.3.3. Selection of parameters' combination to be included in the GIS-based metric

The GIS-based metric was defined through comparison with dioxin concentrations computed by the SIRANE model in Lyon and le Bugey, performed to select the different parameters and their combination to be included in the GIS-based metric. Agreement between quintiles of dioxin concentrations from modeling and quintiles of GIS-based metric estimations were calculated with weighted kappa coefficients and their 95% CI that assigned less importance to discrepancy between adjacent quintiles and higher importance to other discrepancies (Cordioli et al. 2013). Weighted kappa coefficients greater than 0.60 represented a substantial agreement; weighted kappa coefficients greater than 0.80 represented an almost perfect agreement (Viera et al. 2005).

3. Results

3.1. Inventory and characterization of dioxin emitting sources

From 1990 to 2008, N=2,626 dioxin emitting source operated in France, including N=235 in the Rhône-Alpes region, that corresponded to 2,999 distinct operation periods named 'source-periods', (N=302 in the Rhône-Alpes region) according to technological changes. In the Rhône-Alpes region, the predominant sector was waste incineration (N=131 source-periods, 43.4%), followed by production of mineral (N=90, 29.8%), heat and power generation (N=40, 13.3%), metal production (N=27, 8.9%), crematoria (N=13, 4.3%) and chemicals and consumer goods (N=1, 0.3%) (Table 1).

Table 1. Frequency (n) and percentage (%) of industrial sources and source-periods inventoried in the Rhône-Alpes region between 1990 and 2008, and mean and standard deviation (SD) of their annual dioxin emission intensity (in g-TEQ/year). GEO3N project.

Toolkit categories (Toolkit code)	Inventory		Annual dioxin emission intensity (in g-TEQ/year)
	Industrial sources, n (%)	Source-periods, n (%)	Source-periods, mean (SD)
Waste incineration (1.)	92 (39.2)	131 (43.4)	4.0E0 (7.7E0)
Metal production (2.)	21 (8.9)	27 (8.9)	1.7E-1 (2.9E-1)
Heat and power production (3.)	36 (15.3)	40 (13.3)	9.1E-3 (1.7E-2)
Production of mineral (4.)	72 (30.6)	90 (29.8)	3.5E-2 (1.2E-1)
Chemical and consumer goods (7.)	1 (0.4)	1 (0.3)	7.1E-3 ^a
Crematoria (8.b.)	13 (5.5)	13 (4.3)	7.2E-3 (3.5E-3)
Total	235 (100.0)	302 (100.0)	1.8E0 (5.4E0)

^a: one value for annual dioxin emission intensity (no mean, no SD)

In the Rhône-Alpes region, the average annual dioxin emission intensity was estimated at 1.8 ± 5.4 g-TEQ/year between 1990 and 2008. Although, industrial sources from the waste incineration sector released in average 4.0 ± 7.7 g-TEQ/year, the average annual dioxin emission intensities did not exceed 0.5 g-TEQ/year for most industrial sectors (Table 1).

The sum of annual dioxin emission intensities estimated in the Rhône-Alpes region has sharply decreased between 1990 and 2008, with a first major drop in emissions observed in the early 2000s (from 4.9E3 g-TEQ/year to 2.0E2 g-TEQ/year in 2001 and to 3.4E1 g-TEQ/year in 2003) and confirmed by another drop after 2005 to 6.9 g-TEQ/year. Between 2006 and 2008, the sum of annual dioxin emission intensities stabilized at approximately 5.0 g-TEQ/year. The global decrease in annual dioxin emission intensities between 1990 and 2008 was approximately 99.1% (Figure 9).

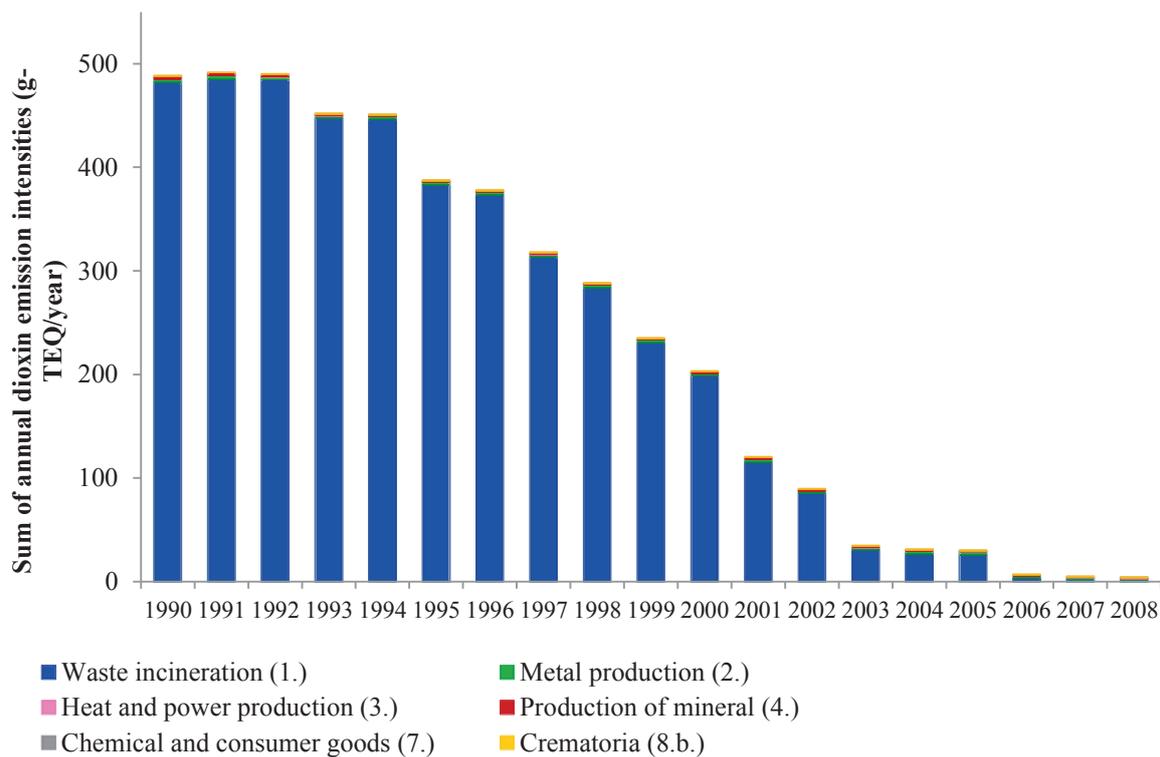


Figure 9. Evolution of the sum of annual dioxin emission intensities estimated between 1990 and 2008, per activity sectors, among the source-periods inventoried within the Rhône-Alpes region. GEO3N project. France, 1990-2008.

3.2. Geocoding

3.2.1. Geocoding of the participants' residential history

Results of the geocoding of the residential addresses are described in Article #2. The automatic geocoding by the ArcGIS software (ArcGIS Locator, ESRI, Redlands, CA, USA) showed better agreement with manual geocoding. Moreover, this method allowed better control at all steps of the geocoding process and was less time consuming. With this method, 61.4% of the addresses were geocoded to the address, 67.1% of which with a positional error of less than 25 m. Only 14.7% of addresses required manual checking. Automatic geocoding with ArcGIS was therefore applied to geocode the residential addresses on the whole French territory.

3.2.2. Geolocation of the dioxin emitting sources

Among the 302 source-periods inventoried in the Rhône-Alpes region between 1990 and 2008, 72% were positioned at the stack, 18% at the centroid of the building, 9% at the centroid of the plot and 1% was located at the town hall. The same method was used for the localization of all industrial sources inventoried in France.

3.3. GIS-based airborne dioxin exposure assessment

3.3.1. Literature review

From the 2,355 articles identified from the literature search, 2,320 focused on air pollution assessment in methodological and epidemiological studies, including 40 on dioxin exposure; 35 articles used a dispersion modeling. Among them, 80 were selected after review of titles and abstracts.

These 80 articles were independently assessed by four investigators (TC, AD, EF and GH), and 7 finally were selected for identification of relevant parameters to be included in the GIS (Gulliver and Briggs 2011; Hoek et al. 2001; Pronk et al. 2013; Vienneau et al. 2009; White et al. 2009; Yu 2006; Zou et al. 2009); the 7 selected publications are presented in Appendice 2. One article focused on dioxin exposure assessment (Pronk et al. 2013), while five others globally focused on air pollution assessment (Hoek et al. 2001; Vienneau et al. 2009; White et al. 2009; Yu 2006) including a review of air pollution exposure methods (Zou et al. 2009); one article used an atmospheric dispersion modeling to assess air pollution exposure (Gulliver and Briggs 2011). Several parameters were identified from the literature review. Residence-to-source distance was computed in 5 studies, residential history was recorded in 3 (Hoek et al. 2001; Pronk et al. 2013; Vienneau et al. 2009; White et al. 2009; Yu 2006) and several distance declines were modeled ($1/d$; $1/d^2$ and $1/d^3$). Meteorological data were used in four studies, e.g. wind speed and wind direction, and wind direction was divided by 45° and 90° segments (Gulliver and Briggs 2011; Vienneau et al. 2009; White et al. 2009; Yu 2006). Different sizes around industrial sites regarding the impact of air pollutant emissions were defined: 11 km, 3 km and 50 m for proximity to roads (Vienneau et al. 2009; Yu 2006). Additional information on stack height was used with the dispersion modeling (Gulliver and Briggs 2011). The literature review allowed the identification of the several potentially influencing airborne dioxin exposure: residence-to-source distance and distance decline; wind direction and wind speed; stack height and exhaust smoke velocity; and an impact zone around industrial sources.

3.3.2. GIS parameters

Residence-to-source distance and the prevailing wind direction were relevant parameters and were integrated in the GIS as follow.

Residence-to-source distance

Distances between residential addresses and sources were calculated within a 10 km around the industrial sources, and computed into a residence-to-source distance matrix. Outside 10 km, dioxin releases from industrial sources were considered negligible and women were considered not exposed. This threshold was confirmed by simulations results, which showed that the impact of the most dioxin emitting source did not exceed 10 km.

Prevailing wind direction

Hourly data on prevailing wind direction were provided by *Météo France*® and averaged over the year for each year between 1990 and 2008. Several “symposium areas” (N=727) were defined by *Météo France*® as homogeneous weather zones sharing the same local meteorology and topography (plains, mountains, valleys, hills...) and cut up within the French territory. Each industrial source was assigned to a symposium area in order to be assigned to a meteorological station. If more than one station was present in a symposium area, the meteorological station with the lengthiest temporal coverage of the study period and most detailed information on direction and speed wind, precipitation and temperature were available was selected. If no meteorological station was present in a specific symposium area, a meteorological station situated in a neighbour symposium area and sharing the same topographical properties was selected. In the Rhône-Alpes region, the 232 sources were assigned to 56 distinct meteorological stations situated in 63 symposium areas. Equal prevailing wind direction segments were determined around each industrial source from 1990 to 2008 and assigned an annual prevailing wind frequency.

3.3.3. Parameters' combination to be included in the GIS-based metric

Parameters tested

Several combinations of parameters were tested, including the Toolkit-based annual dioxin emission intensity of the industrial sources; an impact zone, named “buffer”, of dioxin emissions around each industrial source (3; 5 and 10 km); a residence-to-source distance decline pattern ($1/d$; $1/d^{1.5}$; $1/d^2$; $1/d^3$; e^{-d} ; e^{-d}/d and $e^{-d}*d$); a prevailing wind frequency with segment sizes of 90° , 30° , 10° and 10° with weighted contribution of adjacent segments (Figure 10); wind speed (w); stack height ($1/h$; $1/\sqrt{h}$ and $1/h^2$) and exhaust smoke velocity ($1/(h*v)$; $1/(\sqrt{h*v})$ and $1/(h^2*v)$).

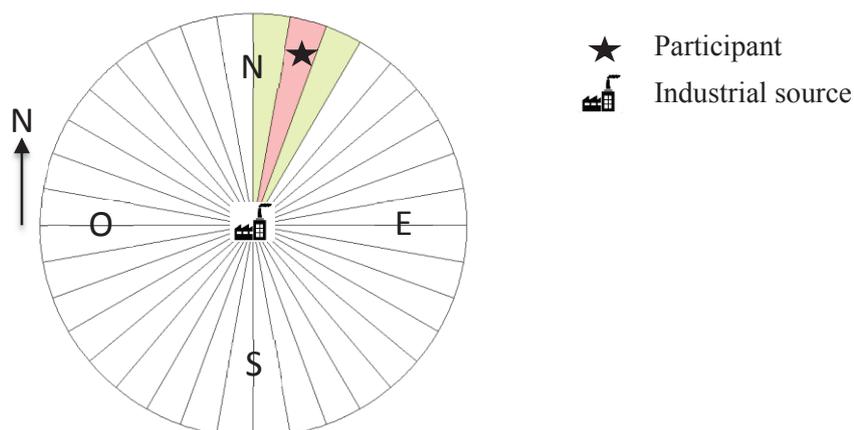


Figure 10. Pattern of prevailing wind direction according to 10° segments (2-times weighted) with contribution of adjacent segments. GEO3N project. 1990-2008.

Weighted kappa coefficients

Results of the comparison between the dispersion modeling data and GIS-based metric estimations according to the distinct parameters' combinations are presented in Appendice 3. In Lyon, the highest weighted kappa coefficients were observed for a residence-to-source distance decline of $1/d^2$ and a prevailing wind direction by 10° segments (2-times weighted) with contribution of adjacent segments. Weighted kappa coefficients ranged from 0.71 (0.67-0.76) to 0.84 (0.79-0.88) (Table 2). This parameters' combination was validated through comparisons with dioxin concentration modeling in Le Bugey, with weighted kappa coefficients ranging from 0.73 (0.66-0.79) to 0.82 (0.76-0.87) (Table 2).

Further kappa statistics were computed, including wind speed, stack height and exhaust smoke velocity in the GIS-based metric. Inclusion of the wind speed led to weighted kappa coefficients ranging from 0.59 (0.52-0.66) to 0.71 (0.61-0.81) in Lyon and from 0.58 (0.49-0.66) to 0.76 (0.70-0.82) in Le Bugey. When including both stack height and exhaust smoke velocity in the GIS-based metric, weighted kappa coefficients were not further improved: weighted kappa coefficients ranged from 0.57 (0.51-0.64) to 0.82 (0.77-0.87) in Lyon and from 0.55 (0.46-0.64) to 0.81 (0.75-0.86) in Le Bugey (Table 6 and Table 7 in Appendice 3).

Table 2. Weighted kappa coefficients (wk) and 95% CI observed in Lyon and Le Bugey for 1996, 2002 and 2008 with a GIS-based metric including a distance decline of $1/d^2$ and a prevailing wind direction divided by 10° segments (2-times weighted) with contribution of adjacent segments.

	1996	2002	2008
	wk (95% CI)	wk (95% CI)	wk (95% CI)
Lyon	0.71 (0.67-0.76)	0.84 (0.79-0.88)	0.81 (0.72-0.89)
Le Bugey	0.79 (0.73-0.85)	0.82 (0.76-0.87)	0.73 (0.66-0.79)

GIS-based metric for airborne dioxin exposure assessment

The selected parameters were integrated in the final GIS-based metric using the following formula:

$$\text{GIS – based metric} = \sum_{j=1}^J \sum_{i=1}^I \left(EI \times t \times \frac{1}{d^2} \times F \right)$$

Where j is the place of residence ($j=1, \dots, J$), i is the industrial source ($i=1, \dots, I$), EI is the annual dioxin emission intensity (in g-TEQ/year), t is the exposure duration (in year), d is the residence-to-source distance (in m) and F is the prevailing wind frequency. The GIS-based airborne dioxin exposure metric was expressed in $\mu\text{g-TEQ}/\text{m}^2$ and was computed for each calendar year from 1990 to 2008.

Wind speed, stack height, exhaust smoke velocity were not considered in the formula because they did not improved agreement with modeling. Moreover, collecting retrospectively stack height and exhaust smoke velocity for each industrial source was difficult and the information was not available for all source-periods.

4. Article #2

Accuracy of two geocoding methods for Geographic Information System based exposure assessment in epidemiological studies

Elodie Faure¹, Aurélie MN Danjou^{1,2}, Charlotte Carretero¹, Françoise Clavel-Chapelon^{3,4,5}, Marie-Christine Boutron-Ruault^{3,4,5}, Laure Dossus^{3,4} and Béatrice Fervers^{1,2}

¹ Département Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France

² Université de Lyon, Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69100 Villeurbanne, France

³ Inserm, Center for Research in Epidemiology and Population Health (CESP), U1018, Lifestyle, genes and health team, 114 rue Edouard-Vaillant, 94805 Villejuif Cedex, France

⁴ Université Paris-Sud, UMRS 1018, Villejuif, France

⁵ Gustave Roussy, F-94805 Villejuif, France

Reference:

Faure E, Danjou AMN, Carretero C, Clavel-Chapelon F, Dossus L, Fervers B. ‘Accuracy of two geocoding methods for Geographic Information System based exposure assessment in epidemiological studies’. *Submitted to Environmental Health on the 04/10/2016.*

Environmental Health

Accuracy of two geocoding methods for Geographic Information System-based exposure assessment in epidemiological studies

--Manuscript Draft--

Manuscript Number:													
Full Title:	Accuracy of two geocoding methods for Geographic Information System-based exposure assessment in epidemiological studies												
Article Type:	Methodology												
Section/Category:	Environmental Epidemiology												
Funding Information:	<table border="1"> <tr> <td>French League against Cancer</td> <td>Research director Françoise CLAVEL-CHAPELON</td> </tr> <tr> <td>Institut Gustave-Roussy</td> <td>Research director Françoise CLAVEL-CHAPELON</td> </tr> <tr> <td>Institut National de la Santé et de la Recherche Médicale</td> <td>Research director Françoise CLAVEL-CHAPELON</td> </tr> <tr> <td>Auvergne Rhône-Alpes Region</td> <td>Associate Professor Béatrice FERVERS</td> </tr> <tr> <td>French Environment and Energy Management Agency - ADEME (CIDE 2012-1)</td> <td>Associate Professor Béatrice FERVERS</td> </tr> <tr> <td>Mutuelle Générale de l'Education Nationale (FR)</td> <td>Research director Françoise CLAVEL-CHAPELON</td> </tr> </table>	French League against Cancer	Research director Françoise CLAVEL-CHAPELON	Institut Gustave-Roussy	Research director Françoise CLAVEL-CHAPELON	Institut National de la Santé et de la Recherche Médicale	Research director Françoise CLAVEL-CHAPELON	Auvergne Rhône-Alpes Region	Associate Professor Béatrice FERVERS	French Environment and Energy Management Agency - ADEME (CIDE 2012-1)	Associate Professor Béatrice FERVERS	Mutuelle Générale de l'Education Nationale (FR)	Research director Françoise CLAVEL-CHAPELON
French League against Cancer	Research director Françoise CLAVEL-CHAPELON												
Institut Gustave-Roussy	Research director Françoise CLAVEL-CHAPELON												
Institut National de la Santé et de la Recherche Médicale	Research director Françoise CLAVEL-CHAPELON												
Auvergne Rhône-Alpes Region	Associate Professor Béatrice FERVERS												
French Environment and Energy Management Agency - ADEME (CIDE 2012-1)	Associate Professor Béatrice FERVERS												
Mutuelle Générale de l'Education Nationale (FR)	Research director Françoise CLAVEL-CHAPELON												
Abstract:	<p>Background: Environmental exposure assessment based on Geographic Information Systems (GIS) and study participants' residential proximity to environmental exposure sources relies on the positional accuracy of subjects' residences to avoid misclassification bias. Our study compared the positional accuracy of two automatic geocoding methods to a manual reference method.</p> <p>Methods: We geocoded 4,247 address records representing the residential history (1990-2008) of 1,685 women from the French national E3N cohort living in the Rhône-Alpes region. We compared two automatic geocoding methods, a free-online geocoding service (method A) and an in-house geocoder (method B), to a reference layer created by manually relocating addresses from method A (method R). For each automatic geocoding method, positional accuracy levels were compared according to the urban/rural status of addresses and time periods (1990-2000, 2001-2008), using Chi Square tests. Kappa statistics were performed to assess agreement of positional accuracy of both methods A and B with the reference method, overall, by time periods and by urban/rural status of addresses.</p> <p>Results: Respectively 81.4% and 84.4% of addresses were geocoded to the exact address (65.1% and 61.4%) or to the street segment (16.3% and 23.0%) with methods A and B. In the reference layer, geocoding accuracy was higher in urban areas compared to rural areas (74.4% vs. 10.5% addresses geocoded to the address or interpolated address level, $p < 0.0001$); no difference was observed according to the period of residence. Compared to the reference method, median positional errors were 0.0 m (IQR=0.0-37.2 m) and 26.5 m (8.0-134.8 m), with positional errors < 100 m for 82.5% and 71.3% of addresses, for method A and method B respectively. Positional agreement of method A and method B with method R was 'substantial' for both methods, with kappa coefficients of 0.60 and 0.61 for methods A and B, respectively.</p> <p>Conclusion: Our study demonstrates the feasibility of geocoding residential addresses in epidemiological studies not initially recorded for environmental exposure assessment, both for recent addresses and residence locations more than 20 years ago. Accuracy of the two automatic geocoding methods was comparable. The in-house method (B) allowed a better control of the geocoding process and was less time consuming.</p>												
Corresponding Author:	Marie-Christine BOUTRON-RUAULT, MD, PhD Gustave Roussy Institute VILLEJUIF Cedex, FRANCE												

Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Gustave Roussy Institute
Corresponding Author's Secondary Institution:	
First Author:	Elodie FAURE, Master degree
First Author Secondary Information:	
Order of Authors:	Elodie FAURE, Master degree
	Aur�lie Marcelle, Nicole DANJOU, MPH, Phd student
	Fran�oise CLAVEL-CHAPELON, PhD
	Marie-Christine BOUTRON-RUAULT, MD, PhD
	Laure DOSSUS, PhD
	B�atrice FERVERS, MD, PhD,
Order of Authors Secondary Information:	
Opposed Reviewers:	John R Nuckols, Emeritus Professor Colorado State University jnuckols@colostate.edu Conflicts of interest
	Mary H WARD, PhD National Cancer Institute wardm@mail.nih.gov Conflict of interests

Author's manuscript (submitted)

Accuracy of two geocoding methods for Geographic Information System-based exposure assessment in epidemiological studies

5 **Elodie Faure**¹, Aurélie MN Danjou^{1,2}, Françoise Clavel-Chapelon^{3,4,5}, Marie-Christine
Boutron-Ruault^{3,4,5}, Laure Dossus^{3,4}, Béatrice Fervers^{1,2}

¹ Cancer and Environment Department, Centre Léon Bérard, 28 rue Laennec, 69373 Lyon
Cedex 08, France

10 ² Claude Bernard Lyon 1 University, 43 Boulevard du 11 Novembre 1918, 69100
Villeurbanne, France

³ Inserm, Centre for research in Epidemiology and Population Health (CESP), U1018, Team
"Generations for Health", 94805 Villejuif, France

⁴ Paris Sud University, UMRS 1018, 94805 Villejuif, France

⁵ Gustave Roussy Institute, 94805 Villejuif, France

15

E-mail addresses

Elodie.FAURE@lyon.unicancer.fr, Aurelie.DANJOU@lyon.unicancer.fr,
Francoise.CLAVEL@gustaveroussy.fr, Marie-Christine.BOUTRON@gustaveroussy.fr,
DossusL@iarc.fr, Beatrice.FERVERS@lyon.unicancer.fr

20

Corresponding author

Marie-Christine BOUTRON-RUAULT - Marie-christine.BOUTRON@gustaveroussy.fr
INSERM U1018 – EMT, Institut Gustave Roussy, 114 rue Edouard Vaillant, 94805
VILLEJUIF Cedex, FRANCE

25 Phone: ++331 42 11 64 66

Fax: +331 42 11 40 00

Manuscript submitted to Environmental Health on the 04/10/2016

30 **Abstract**

Background:

Environmental exposure assessment based on Geographic Information Systems (GIS) and study participants' residential proximity to environmental exposure sources relies on the
35 positional accuracy of subjects' residences to avoid misclassification bias. Our study compared the positional accuracy of two automatic geocoding methods to a manual reference method.

Methods:

We geocoded 4,247 address records representing the residential history (1990-2008) of 1,685
40 women from the French national E3N cohort living in the Rhône-Alpes region. We compared two automatic geocoding methods, a free-online geocoding service (method A) and an in-house geocoder (method B), to a reference layer created by manually relocating addresses from method A (method R). For each automatic geocoding method, positional accuracy levels were compared according to the urban/rural status of addresses and time periods (1990-2000,
45 2001-2008), using Chi Square tests. Kappa statistics were performed to assess agreement of positional accuracy of both methods A and B with the reference method, overall, by time periods and by urban/rural status of addresses.

Results:

Respectively 81.4% and 84.4% of addresses were geocoded to the exact address (65.1% and
50 61.4%) or to the street segment (16.3% and 23.0%) with methods A and B. In the reference layer, geocoding accuracy was higher in urban areas compared to rural areas (74.4% vs. 10.5% addresses geocoded to the address or interpolated address level, $p < 0.0001$); no difference was observed according to the period of residence. Compared to the reference method, median positional errors were 0.0 m (IQR=0.0-37.2 m) and 26.5 m (8.0-134.8 m),
55 with positional errors < 100 m for 82.5% and 71.3% of addresses, for method A and method B respectively. Positional agreement of method A and method B with method R was 'substantial' for both methods, with kappa coefficients of 0.60 and 0.61 for methods A and B, respectively.

Conclusion:

60 Our study demonstrates the feasibility of geocoding residential addresses in epidemiological studies not initially recorded for environmental exposure assessment, both for recent addresses and residence locations more than 20 years ago. Accuracy of the two automatic

Author's manuscript (submitted)

geocoding methods was comparable. The in-house method (B) allowed a better control of the geocoding process and was less time consuming.

65

Key words: Geocoding; Geographic Information System – GIS; epidemiology; environmental epidemiology; residential history

Short title: Accuracy of geocoding for GIS exposure assessment

70 **Introduction**

Environmental epidemiology requires reliable assessment of both temporal and spatial components of exposure. In response to these challenges, epidemiological studies are increasingly using residential addresses of study participants and geographic information systems (GIS) to improve characterization of environmental exposures and examine their association with human health risks for a large variety of disease conditions [1]. GIS have, for instance, been used to investigate the relationship between environmental exposures and risk of breast cancer [2–4], leukemia [5–7], Parkinson's diseases [8,9], adverse birth outcomes [10,11], or respiratory health [12–15]. GIS-based exposure assessment using residential proximity to the environmental exposure source (e.g. farmland treated with pesticides, industrial facilities or traffic roads) as an exposure surrogate relies on the positional accuracy of the subjects' residences to avoid exposure misclassification [16]. There is increasing use of existing prospective cohorts for investigating environmental causes of diseases, although most of them had not been initially designed for environmental exposure assessment [17,18]. While the strength of using existing cohorts relies on the prospective data collection at the individual level over many years allowing to adjust for individual disease risk factors, the subjects' postal addresses have rarely been collected to be geocoded (i.e. to be converted into precise geographic coordinates) for their use in GIS. This may result in poor positional accuracy of subjects' addresses and may represent an important source of misclassification and imprecision in environmental exposure assessment [13,16,19–24].

The process of geocoding and assigning geographic coordinates (latitude and longitude) to the study subject's residential addresses is one of the first steps in GIS-based epidemiological studies [20,24,25]. The quality of geocoding depends on the completeness and the level of positional accuracy of located addresses. Completeness is the proportion of addresses that can be geocoded and depends on the quality of the collected data on addresses. The positional accuracy reflects the level of proximity of geocoded objects to their true location [26,27]. Geocoding residential addresses can be performed using three methods. A first method consists in using online geocoding services to obtain subjects' coordinates or to create online maps with subjects' residence locations. These free services are available on the Internet and do not require specific expertise in geocoding [21]. A second approach consists in using a commercial service that can handle all steps of geocoding from the spell checking of addresses to their map location [11,13,24,28]. The third method is the use of an in-house method of geocoding [7,21,24,29] where the geocoding process is handled by the research

Author's manuscript (submitted)

105 team using commercially available GIS software equipped with a geocoding tool and a
reference street database. In Europe, and particularly in France, there is a lack of studies
comparing accuracy of geocoding between different geocoding tools, as well as according to
characteristics of residential locations and date of residence.

110 Several American and European studies have evaluated the accuracy of different
geocoding methods and of their reference network database in comparison to field location
using Global Positioning System (GPS) [13,20,26,30] and manual location based on aerial-
photography [27,31]. These studies have raised awareness on the divergence of geocoding
accuracy between methods, with variations in median positional errors ranging from 25 m to
201 m. Also, accuracy levels of geocoding may vary according to the urban or rural status of
the subjects' residential location [20,24,31,32]. Furthermore, studies investigating differences
115 in geocoding accuracy of residential addresses by date of residence yielded inconsistent
results [20,32].

Our study aimed at comparing the accuracy of two automatic geocoding methods, an
online method and an in-house method, with a manual method of geocoding used as the
reference, in a French national prospective cohort initiated in 1990. Our study further assessed
120 the geocoding accuracy according to urban and rural status of the addresses and the period of
residence. The study was performed in order to subsequently use the most suitable method for
geocoding of subjects' residences to assess environmental exposure in a case-control study
nested within the same prospective cohort.

125 **Methods**

Study Population

Our analysis involved study subjects from a nested case control study (the Geo3N research project) including 5,455 breast cancer cases and 5,455 matched controls that aimed at
130 analyzing the association between environmental dioxin exposure and breast cancer risk in the E3N (*Etude Epidémiologique des Femmes de la Mutuelle Générale de l'Education Nationale*) cohort [33]. E3N is an ongoing prospective French cohort study of 98,995 women investigating female cancer risk factors. E3N is the French component of the European Prospective Investigation into Cancer and Nutrition [34]. E3N participants were enrolled in
135 1990 at the age of 40 to 65 years old and were members of a national teachers' health insurance. Subjects are followed up by self-administered questionnaires every 2 to 3 years. E3N was approved by the French commission for Data Protection and Privacy. For the present analysis, we selected 1,730 subjects from the nested case control study. All selected subjects lived in the Rhône-Alpes region at recruitment in the E3N cohort. The Rhône-Alpes
140 region covers a territory of 43,196 km² with over 6 million inhabitants and presents a broad diversity of territories with rural, mountainous, and highly urbanized areas. Residential addresses of study participants were collected through the baseline and four follow-up questionnaires, sent in 1990, 1997, 2000, 2002, and 2005 respectively.

145 ***Data cleaning***

To improve standardization and quality of geocoding, all subjects' addresses were verified manually for the spelling of street and municipality names, using free online databases referencing French postal codes (e.g. www.codespostaux.com/, www.pagesjaunes.fr/pagesblanches/). We also completed address fields of subjects (i.e.
150 missing or incomplete postal code, municipality name, street name and street number) by matching with data from previous and subsequent questionnaires for similar residential location (e.g. same street name and same city name). After exclusion of 45 subjects with missing address, postal code, or municipality name, 1,685 subjects corresponding to a total of 4,247 addresses consecutively collected at each questionnaire between 1990 and 2008 were
155 included in the analysis.

Geocoding methods

For each assessed method, dots representing addresses were located along the street at the entrance of the building. A trained technician geocoded all addresses blinded to the case-control status of the subjects.

160 *Automatic methods*

The first technique, “method A”, consisted in an automatic method based on a free online geocoding service accessible at www.batchgeocodeur.mapjnz.com/. The reference street network database was based on Google Maps; the total number of addresses stored in the database was not provided by Google. After automatic online geocoding processing, 165 latitude and longitude coordinates in the WGS 84 projection system were exported for each address geocoded as well as accuracy of each location ranking from 0 to 9 (0: not found, 1: country level, 2: region (state and district), 3: county, 4: city, 5: postal code, 6: street segment, 7: intersection of streets, 8: address, 9: point of interest (building names, church...)). In France, levels 2, 3 and 7 did not exist in the administrative division of territories and were 170 therefore not applicable. For geocoded addresses with a precision lower than 6 (street segment level), the spelling of street names and municipalities were checked again manually and corrected if necessary. Revised addresses were then geocoded a second time with the same online geocoder. The database was imported into ArcGIS 10.0 (Environmental System Research Institute- ESRI-Redlands, CA, USA) to create a data layer, and all coordinates were 175 converted into Lambert 93, which is the projection system currently used in France.

The second automatic method, “method B”, consisted in an in-house method based on the BD Adresse for ArcGIS [35,36] and its reference street network database, BD Adresse® (National Geographic Institute, IGN, Saint Mandé, France) that includes 26 million addresses. For each geocoded address, the ArcGIS software provided two indicators to determine the 180 best position: the spelling sensitivity that is the degree to which the spelling variation of a street name is allowed during a search for likely match candidates and the minimum candidate score that is a potential match record requires to be considered a candidate [37]. In this study, we used the following thresholds for the geocoding matching options: 50 for the spelling sensitivity, 80 for the minimum candidate score. For method B, the geocoding accuracy 185 levels ranked from 0 to 7: 0: not found, 1: exact address, 2: interpolated address, 3: street segment, 4: locality, 5: town hall, 6: postal code, 7: city. Interpolated addresses are located based on known positioned addresses along the street. For addresses with several possible matches (N=499), we selected the address with the highest candidate score. Data were projected in Lambert 93.

190 *Reference method*

The reference method in the present study, “method R”, was created by manually relocating addresses located by the online geocoder (www.batchgeocodeur.mapjnz.com) used in method A and the ArcGIS online geocoding service. As it was not feasible, timewise, to check manually all addresses, it was decided to check manually all addresses with an accuracy level equal to or lower than street segment (≤ 6) or all the addresses (accuracy level 8 or 9) deviating more than 50 m from the position obtained with an independent method of geocoding (ArcGIS online geocoding service.). Consequently, all addresses with an accuracy level of 8 (address) or 9 (point of interest: school, hometown, etc.) were geocoded again with the ArcGIS online geocoding service [38]. The street network database of ArcGIS online is based on Navteq ; the number of reference addresses stored in the database was not available from ESRI. For each address geocoded by both method A(batchgeocoder) and ArcGis online, the Euclidian distance between addresses geocoded by the online geocoder and the addresses geocoded by ArcGis Online was computed, using the point to point function in ArcGis 10.0. Addresses with a distance larger than 50 meters were selected for manual checking and verified by a trained technician; dots were relocated to the right location when necessary. For manually checked addresses, relocation and the new accuracy level (town hall, locality, street segment or addresses) were recorded, as well as the information on the database used to determine manually the most accurate location (i.e. Google Maps®, Google Street View®, Geoportail®, Yahoo Map®, Geoportail® from IGN or BD Adresses® from IGN®) [31,39,40]. Addresses were located at the best available location based on the specification of the address itself (e.g., when the street number was missing, the best accuracy would have been to the street segment). When Geo3N addresses did not have exact postal address information (missing street name), the best location was the town hall. “Method R” was used as the reference to compare the accuracy of the two automatic geocoding methods (methods A and B).

Data analyses

To facilitate the comparison of the accuracy of the two automatic geocoding methods with the accuracy of method R, we regrouped the accuracy levels for each method into three categories, *i.e.* city or postal code, street segment, address or point of interest for method A; postal code or town hall, locality or street segment, interpolated address or address for method B; town hall, street segment or locality, address for method R. To assess the geocoding accuracy of each of the two automatic methods, we selected unique addresses among the

Author's manuscript (submitted)

4,247 residential addresses collected consecutively. As the addresses were collected for each
225 questionnaire, if the woman lived at the same address for two consecutive questionnaires,
there were two addresses in the database. Minimal differences could occur in the spelling of
these addresses. To identify unique addresses, we need to know the points that overlap. Thus,
we calculated X and Y coordinates of each automatically geocoded point from methods A and
B, as well as the X and Y coordinates from method R. By matching X and Y coordinates from
230 methods A and R, and methods B and R, we obtained respectively 2,224 and 2,425 pairs of X
and Y coordinates corresponding to unique addresses. For method R, we obtained 2112
unique pairs of X and Y coordinates. We computed two distance matrices, one for method A
and one for method B, by calculating the Euclidian distance between each automatically
geocoded unique address and its corresponding address in method R. We grouped distances
235 into 6 categories (0-25 m, 26-50 m, 51-100 m, 101-400 m, 401-800 m and greater than 800 m)
and the proportion of addresses in each category was computed for methods A and B.

For each address, the urban or rural status was established for the year of residence
using the French national institute for statistics and economic studies (INSEE) data. To
account for changes in status over time, we used the 1990 urban area definition (UAD) for
240 addresses from 1990 to 1995; the 1999 UAD for addresses from 1996 to 2004 and the 2010
UAD for addresses from 2005 to 2008. For methods A and B, levels of accuracy were
compared between urban and rural areas and according to two periods (1990-2000 and 2001-
2008) using Chi-Square tests. All p-values were two-sided and the significance level was set
to 0.05. Cohen's kappa coefficients were calculated to assess the agreement between the
245 accuracies of methods A and B and the accuracy of method R [41]. The kappa coefficients
were also calculated by time periods (1990-2000 and 2000-2008) and urban/rural status. We
used the SAS statistical software version 9.4 (SAS Institute Inc., Cary, North Carolina) for
data analysis.

250 **Results**

During the data cleaning step, the spelling of 235 addresses (5.5%) and 991 city names (23.3%) was corrected. Over the study period (1990-2008), more than 80% of the study population lived in urban areas and 75% of the study subjects remained at the same address throughout follow-up.

Based on 2,112 unique pairs of X and Y coordinates of addresses coded by method R, 723 (34.2%) addresses with an accuracy level lower or equal to 6 were checked manually and 329 (45.5%) of the latter were relocated. Also, 1,389 (65.8%) addresses with accuracy levels of 8 and 9 were re-geocoded with ArcGis online and 203 (14.6%) of the latter were relocated manually. Overall, with this reference layer, 63.2% of residence addresses were located to the address, 29.2% to the street segment or to the locality and 7.6% to the town hall. In the reference layer, 74.4% and 10.5% of rural addresses were geocoded at the address or interpolated address levels ($p < 0.0001$). The level of accuracy did not vary according to the time period of residence ($p = 0.67$): 62.1% and 63.3% of addresses were located to the address level for the periods 1990-2000 and 2001-2008, respectively.

The positional errors of addresses located by method A and method B compared with method R are presented in table 1. With method A, 405 (18.2%) addresses had a level of accuracy to the city or to the postal code, 363 (16.3%) to the street segment, 1,448 (65.1%) to the address or to the point of interest; 8 addresses (0.4%) could not be geocoded. Among addresses geocoded to the street segment, and address or point of interest, 226 (62.3%) and 1,241 (85.7%) respectively had a positional error of less than 25 m when compared with the layer generated by method R. For addresses geocoded to the city level or postal code, 160 (39.5%) had a positional error lower than 25 m and 146 (36.1%) above 400 m. Seventeen addresses (0.8%) had a positional error of over 30 km. The latter were incomplete and presented a wrong spelling of the municipality. Using method B, 1,490 (61.4%) of the addresses were geocoded to the point address or interpolated address level, 558 (23.0%) to the street segment or locality, and 377 (15.5%) to the town hall or postal code. One thousand (67.1%) addresses were located to the address or to interpolated address with a positional error of less than 25 m, as well as 53 (14.1%) were located at the postal code or town hall level, and 132 (23.7%) were located at the street segment. Addresses with the highest positional error (> 400 m) compared with method R were geocoded to the town hall or the postal code ($n = 244$, 64.8%). Overall, 14.7% of addresses required manual checking with

Author's manuscript (submitted)

method B compared with 30.7% with method A, resulting in less geocoding time spent by the technician for method B.

285 The concordance with method R was assessed separately for each of the two automatic methods (table 2). Overall, Kappa coefficients were 0.60 between methods A and R and 0.61 between methods B and R. For addresses located in urban areas agreements of 0.56 and 0.52 were found respectively between methods A and R, and methods B and R, while in rural areas, agreements with method R were 0.39 and 0.54 for methods A and B respectively.

290 Agreement with method R was 0.61 and 0.60 respectively for methods A and B for the period 1990-1999 and 0.56 and 0.70 respectively for methods A and B for the period 2000-2008.

Discussion

295 In the present study, we compared the accuracy of two automatic geocoding methods, overall,
and according to urban or rural status of addresses and to the time period of residence (from
1990 to 2008). Compared with the reference method, the two methods of geocoding gave
similar results in terms of general accuracy, with more than 60% of addresses geocoded to the
exact address level, and more than 15% to the street segment level. Accuracy was higher in
300 urban areas than in rural areas, while no difference was observed according to the period of
residence. Compared with method A, method B allowed more control at all steps of the
geocoding process and was less time consuming, in particular regarding the manual checking

Based on the Euclidian distance to the address located by method R, 82.5% of
addresses geocoded with method A and 71.3% of addresses with method B had a positional
305 error lower than 100 m. This difference can be explained by the use of the same initial
automatic geocoding (online geocoding) for method A and method R. Overall, these
proportions are comparable to those from other studies conducted in France and outside
France, with figures of 80.9% to 82.0% of addresses with positional error below 100 m in two
French studies [28,32], and 72.0% to 86.0% in international studies [20,24,30]. While these
310 geocoding errors appear overall modest in magnitude, Ganguly et al. showed that positional
errors exceeding 100 m may alter exposure estimates, in particular for exposures with
important spatial gradients, such as traffic-related air pollution [16]. These observations stress
the importance of conducting sensitivity analyses to examine the potential impact of
positional errors on exposure estimates.

315 Our findings are consistent with other studies showing a more precise and accurate
geocoding for addresses located in urban areas compared to rural areas [20,24,31,32]. In these
studies, the median values of the positional error ranked from 31 m to 56 m in urban areas and
from 45 m to 212 m in rural areas where addresses lack frequently street number and are often
limited to the name of the hamlet. Three studies (two US and one French) have geocoded
320 historical addresses covering periods ranking respectively from 1948 to 2000, 1930 to 2000,
and 1960 to 2001 [2,20,32]. In agreement with the present study, two of them did not observe
major variations in the positional accuracy according to the timing of addresses [20,32],
whereas Brody et al. , a US study, reported a better positional accuracy for recent addresses,
with 37% of addresses located to the address level in 1930 vs. 62% for 1970-1980 and 97%
325 for 1990-2001 [2]. Kappa coefficients showed overall good agreement with the reference
method for the two automatic methods. However, for rural addresses, agreement was higher

Author's manuscript (submitted)

with method B compared with method A (0.54 vs. 0.39, respectively). Also, recent addresses (from 2000-2008) showed a higher Kappa coefficient for method B compared with method A (0.70 vs. 0.56).

330 The strengths of our study include the large number of addresses (n=4,247) and study
subjects (n=1,685) for whom residential addresses had been prospectively recorded over a 19-
year study period (1900-2008). Moreover, we were able to classify addresses according to the
rural or urban status of the area of residence. Our study is one of the first to investigate the
geocoding of subjects from a national prospective cohort and offer both a spatial and temporal
335 analysis on the quality of geocoding using different tools. In addition, the manual checking of
the correct location of addresses was done for a large number of addresses and based on aerial
images. However, our study has several limitations. First, E3N addresses were not initially
designed to be geocoded and this could have affected positional accuracy. However, the
findings on addresses located with a positional error lower than 100 m were consistent with
340 another French study in which addresses were recorded to be geocoded [32]. Also, to take
into account potential errors in the spelling of addresses not collected to be geocoded we used
for method B a threshold of 50 for spelling sensitivity. A higher threshold, as used by some
authors would have allowed only minor spelling variations of addresses and restrict the
candidates to exact matches [21,24,36,40]. As E3N participants are mostly teachers, some of
345 them indicated only the name of their school (n=49) or workplace (n=6) in the address field;
because these names are not available as such in reference databases, those could not be
automatically geocoded and this may have had a minor impact on the global accuracy of both
automatic geocoding methods. Second, because of the large number of addresses in our study
and the covered territory, feasibility of using field GPS location to validate the true location of
350 all addresses, as performed by others was limited [13,20,24,30]. However, the use of aerial
photography via Google Maps, Geoportail and Google Street View® to manually check all
addresses initially geocoded with a low level of accuracy, allowed us to be confident in the
precision of the address location in our reference layer (method R). Third, one weakness of
methods A and R is the impossibility, despite repeatedly contacting Google and ESRI France,
355 to obtain the number of addresses recorded in their reference database.

Furthermore, geocoding of addresses in environmental epidemiology using external
services or free online devices, such as the batch online geocoding, raises privacy and ethical
considerations [42]. Since addresses may allow the personal identification of the study
subjects, their transfer to third parties may breach participants' confidentiality and anonymity,
360 even after removal of any sensitive information, and in particular in defined geographic areas

Author's manuscript (submitted)

with small numbers of study subjects. In-house geocoding generally allows a better control of any type of unauthorized access to sensitive information.

Conclusion

365

Our study demonstrated the feasibility of geocoding addresses in epidemiological studies not initially designed to be used for environmental exposure assessment purposes, for both recent addresses and residence locations dated from more than 20 years. Furthermore, our results showed no major difference in final geocoding accuracy between the two automatic

370

geocoding methods, compared with the manual reference method. Overall, more addresses showed a positional error lower than 100 m with method A, while the Kappa coefficients showed higher agreement with the reference method for method B, for both rural areas and the 2000-2008 period. Also, this in-house method allowed a better control at all steps of the geocoding process and was less time consuming. Future epidemiological studies should

375

prospectively record residential addresses in a way that would improve geocoding for environmental exposure assessment. Finally, knowing the accuracy of the geocoding tool used in the context of environmental exposure assessment will help to limit misclassification bias due to positional errors. Epidemiological studies should be able to report their street network reference database and the accuracy of their geocoding method.

380

Author's manuscript (submitted)

List of abbreviations

ESRI :Environmental Systems Research Institute

385 GIS :Geographical Information System

GPS : Global Positioning System

IGN : National Geographic Institute

INSEE :French national institute for statistics and economic studies

UAD :urban area definition

390 WGS : World Geodetic System

Declarations

395 *Ethics approval and consent to participate*

The E3N cohort study received ethical approval from the French National Commission for Computerized Data and Individual Freedom (Commission Nationale de l'Informatique et des Libertés, CNIL), and all women in the study provided signed informed consent for participation into the study. This approval covers all questionnaires designed within the
400 framework of the cohort, and no new approval is needed for individual projects.

Consent for publication

Not Applicable

405 *Availability of data and materials*

The datasets generated or analyzed during the current study are available from the corresponding author on reasonable request

Competing interests

410 The authors declare that they have no competing interests.

Funding

The E3N study is financially supported by the French League Against Cancer, the *Mutuelle Générale de l'Education Nationale*, the *Institut Gustave Roussy*, the *Institut National de la*
415 *Santé et de la Recherche Médicale*. The present project received funding from the Auvergne-Rhône-Alpes region and the French Environment and Energy Management Agency.

Authors' contributions

FCC and MCBR have contributed to the conception and design of the E3N study. FCC,
420 MCBR and LD have made contribution to the data acquisition. EF geocoded the addresses and performed the GIS analyses. AMND performed the statistical analyses. EF drafted the manuscript. EF, AMND, LD and BF have interpreted the results. LD, AMND, BF, FCC and MCBR have made contribution to the revision of the manuscript. All authors have read and approved the submitted manuscript.

Author's manuscript (submitted)

425 *Acknowledgment*

We gratefully acknowledge the participants of E3N for providing data. We thank Nicolas FDAOUCH for its involvement in the geocoding of the participants' addresses.

430 **References**

1. Han D, Bonner MR, Nie J, Freudenheim JL. Assessing bias associated with geocoding of historical residence in epidemiology research. *Geospatial Health*. 2013;7:369–374.
2. Brody JG, Aschengrau A, McKelvey W, Rudel RA, Swartz CH, Kennedy T. Breast cancer risk and historical exposure to pesticides from wide-area applications assessed with GIS. *Environ. Health Perspect.* 2004;112:889–97.
3. O'Leary ES, Vena JE, Freudenheim JL, Brasure J. Pesticide exposure and risk of breast cancer: a nested case-control study of residentially stable women living on Long Island. *Environ. Res.* 2004;94:134–44.
4. Reynolds P, Hurley SE, Gunier RB, Yerabati S, Quach T, Hertz A. Residential proximity to agricultural pesticide use and incidence of breast cancer in California, 1988-1997. *Environ. Health Perspect.* 2005;113:993–1000.
5. Gunier RB, Ward MH, Airola M, Bell EM, Colt J, Nishioka M, et al. Determinants of agricultural pesticide concentrations in carpet dust. *Environ. Health Perspect.* 2011;119:970–6.
6. Hoffmann W, Terschüeren C, Heimpel H, Feller A, Butte W, Hostrup O, et al. Population-based research on occupational and environmental factors for leukemia and non-Hodgkin's lymphoma: the Northern Germany Leukemia and Lymphoma Study (NLL). *Am. J. Ind. Med.* 2008;51:246–57.
7. Rull RP, Gunier R, Von Behren J, Hertz A, Crouse V, Buffler PA, et al. Residential proximity to agricultural pesticide applications and childhood acute lymphoblastic leukemia. *Environ. Res.* 2009;109:891–9.
8. Lee P-C, Bordelon Y, Bronstein J, Ritz B. Traumatic brain injury, paraquat exposure, and their relationship to Parkinson disease. *Neurology.* 2012;79:2061–6.
9. Ritz B, Costello S. Geographic model and biomarker-derived measures of pesticide exposure and Parkinson's disease. *Ann. N. Y. Acad. Sci.* 2006;1076:378–87.
10. Boccolini P de MM, Boccolini CS, Meyer A, Chrisman J de R, Guimarães RM, Veríssimo G. Pesticide exposure and low birth weight prevalence in Brazil. *Int. J. Hyg. Environ. Health.* 2013;216:290–4.

Author's manuscript (submitted)

- 460 11. Chevrier C, Serrano T, Lecerf R, Limon G, Petit C, Monfort C, et al. Environmental determinants of the urinary concentrations of herbicides during pregnancy: the PELAGIE mother-child cohort (France). *Environ. Int.* 2014;63:11–8.
12. Gorai AK, Tuluri F, Tchounwou PB. A GIS based approach for assessing the association between air pollution and asthma in New York State, USA. *Int. J. Environ. Res. Public Health.* 2014;11:4845–69.
- 465 13. Jacquemin B, Lepeule J, Boudier A, Arnould C, Benmerad M, Chappaz C, et al. Impact of geocoding methods on associations between long-term exposure to urban air pollution and lung function. *Environ. Health Perspect.* 2013;121:1054–60.
14. Nuvolone D, Barchielli A, Forastiere F, Gruppo collaborativo EpiAir. [Traffic patterns regulations to reduce air pollution in the Italian cities participating in the EpiAir Project]. *Epidemiol. Prev.* 2009;33:103–12.
- 470 15. Ranzi A, Porta D, Badaloni C, Cesaroni G, Lauriola P, Davoli M, et al. Exposure to air pollution and respiratory symptoms during the first 7 years of life in an Italian birth cohort. *Occup. Environ. Med.* 2014;71:430–6.
16. Ganguly R, Batterman S, Isakov V, Snyder M, Breen M, Brakefield-Caldwell W. Effect of geocoding errors on traffic-related air pollutant exposure and concentration estimates. *J. Expo. Sci. Environ. Epidemiol.* 2015;25:490–8.
- 475 17. Beelen R, Hoek G, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, et al. Natural Cause Mortality and Long-Term Exposure to Particle Components: An Analysis of 19 European Cohorts within the Multi-Center ESCAPE Project. *Environ. Health Perspect.* [Internet]. 2015 [cited 2016 Aug 10]; Available from: <http://ehp.niehs.nih.gov/1408095>
- 480 18. Jacquemin B, Siroux V, Sanchez M, Carsin A-E, Schikowski T, Adam M, et al. Ambient air pollution and adult asthma incidence in six European cohorts (ESCAPE). *Environ. Health Perspect.* 2015;123:613–21.
19. Zandbergen PA, Green JW. Error and bias in determining exposure potential of children at school locations using proximity-based GIS techniques. *Environ. Health Perspect.* 2007;115:1363–70.
- 485

Author's manuscript (submitted)

20. Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Freudenheim JL. Positional accuracy of geocoded addresses in epidemiologic research. *Epidemiol. Camb. Mass.* 2003;14:408–12.
21. Duncan DT, Castro MC, Blossom JC, Bennett GG, Gortmaker SL. Evaluation of the
490 positional difference between two common geocoding methods. *Geospatial Health.*
2011;5:265–73.
22. Jacquez GM. A research agenda: does geocoding positional error matter in health GIS studies? *Spat. Spatio-Temporal Epidemiol.* 2012;3:7–16.
23. Kravets N, Hadden WC. The accuracy of address coding and the effects of coding errors.
495 *Health Place.* 2007;13:293–8.
24. Ward MH, Nuckols JR, Giglierano J, Bonner MR, Wolter C, Airola M, et al. Positional accuracy of two methods of geocoding. *Epidemiol. Camb. Mass.* 2005;16:542–7.
25. Rushton G, Armstrong MP, Gittler J, Greene BR, Pavlik CE, West MM, et al. Geocoding in cancer research: a review. *Am. J. Prev. Med.* 2006;30:S16-24.
- 500 26. Ribeiro AI, Olhero A, Teixeira H, Magalhães A, Pina MF. Tools for address georeferencing - limitations and opportunities every public health professional should be aware of. *PloS One.* 2014;9:e114130.
27. Zandbergen PA. Influence of geocoding quality on environmental exposure assessment of children living near high traffic roads. *BMC Public Health.* 2007;7:37.
- 505 28. Sermage-Faure C, Demoury C, Rudant J, Goujon-Bellec S, Guyot-Goubin A, Deschamps F, et al. Childhood leukaemia close to high-voltage power lines--the Geocap study, 2002-2007. *Br. J. Cancer.* 2013;108:1899–906.
29. Jones RR, DellaValle CT, Flory AR, Nordan A, Hoppin JA, Hofmann JN, et al. Accuracy of residential geocoding in the Agricultural Health Study. *Int. J. Health Geogr.* 2014;13:37.
- 510 30. Zhan FB, Brender JD, De Lima I, Suarez L, Langlois PH. Match rate and positional accuracy of two geocoding methods for epidemiologic research. *Ann. Epidemiol.* 2006;16:842–9.
31. Cayo MR, Talbot TO. Positional error in automated geocoding of residential addresses. *Int. J. Health Geogr.* 2003;2:10.

Author's manuscript (submitted)

- 515 32. Béranger R, Blain J, Baudinet C, Faure É, Fléchon A, Boyle H, et al. Tumeurs germinales du testicule et expositions précoces aux pesticides: étude pilote TESTEPERA. *Bull. Cancer (Paris)*. 2014;101:225–35.
33. Clavel-Chapelon F, E3N Study Group. Cohort Profile: The French E3N Cohort Study. *Int. J. Epidemiol.* 2015;44:801–9.
- 520 34. Riboli E, Hunt K, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr.* 2002;5:1113.
35. ESRI France. BD Adresse for ArcGis [Internet]. 2015 [cited 2015 Oct 26]. Available from: http://esrifrance.fr/iso_album/SPECS_BDADRESSE21_v30.pdf
- 525 36. ESRI. ArcGIS Help 10.1 - Geocoding option properties [Internet]. 2015 [cited 2015 Oct 26]. Available from: <http://resources.arcgis.com/en/help/main/10.1/>
37. ESRI. Esri Support GIS Dictionary [Internet]. 2015 [cited 2016 Sep 22]. Available from: <http://support.esri.com/other-resources/gis-dictionary/browse/m>
38. ESRI. Geocode coverage [Internet]. 2015 [cited 2015 Oct 26]. Available from: <https://doc.arcgis.com/en/arcgis-online/reference/geocode-coverage.htm>
- 530 39. IGN IN de l'Information G et F. Le portail des territoires et des citoyens - Géoportail [Internet]. 2015 [cited 2015 Oct 26]. Available from: <http://www.geoportail.gouv.fr/accueil>
40. Schootman M, Sterling DA, Struthers J, Yan Y, Laboube T, Emo B, et al. Positional accuracy and geographic bias of four methods of geocoding in epidemiologic research. *Ann. Epidemiol.* 2007;17:464–70.
- 535 41. Cohen J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* 1960;20:37–46.
42. Mazumdar S, Konings P, Hewett M, Bagheri N, McRae I, Del Fante P. Protecting the privacy of individual general practice patient electronic records for geospatial epidemiology research. *Aust. N. Z. J. Public Health.* 2014;38:548–52.
- 540

Table 1: Comparison of positional errors^a (in meter) of addresses located by two automatic geocoding methods.

Method of geocoding	Level of accuracy (accuracy code)	N addresses (%)	Median distance to the reference method location (IQR) in meters	Min-max, in meters	Distance in meters, N addresses (%)					
					[0-25]]25-50]]50-100]]100-400]]400-800]	>800
Method A ^b (Batch Geocodeur)	Not found (0)	8 (0.4)	5919.5 (815.4-5639076.5)	117.7-5679417.1	-	-	-	2 (25.0)	-	6 (75.0)
	City (4) or Postal code (5)	405 (18.2)	108.2 (0.0-787.3)	0.0-35204.7	160 (39.5)	13 (3.2)	26 (6.4)	60 (14.8)	48 (11.9)	98 (24.2)
	Street segment (6)	363 (16.3)	0.0 (0.0-108.1)	0.0-519630.8	226 (62.3)	18 (5.0)	25 (6.9)	41 (11.3)	10 (2.8)	43 (11.8)
	Address (8) or Point of interest (9)	1448 (65.1)	0.0 (0.0-0.0)	0.0-292249.4	1241 (85.7)	58 (4.0)	66 (4.6)	56 (3.9)	7 (0.5)	20 (1.4)
	Total	2224 (100.0)	0.0 (0.0-37.2)	0.0-5679417.1	1627 (73.2)	89 (4.0)	117 (5.3)	159 (7.1)	65 (2.9)	167 (7.5)
Method B (ArcGis Locator)	Postal code (6) or Town Hall (5)	377 (15.5)	788.9 (114.0-1845.1)	0.3-8545.1	53 (14.1)	11 (2.9)	25 (6.6)	44 (11.7)	56 (14.9)	188 (49.9)
	Locality (4) or Street segment (3)	558 (23.0)	110.3 (26.7-320.6)	0.0-14477.6	132 (23.7)	55 (9.9)	82 (14.7)	174 (31.2)	47 (8.4)	68 (12.2)
	Interpolated address (2) or Address (1)	1490 (61.4)	12.5 (6.0-35.6)	0.0-3951.9	1000 (67.1)	237 (15.9)	133 (8.9)	100 (6.7)	5 (0.3)	15 (1.0)
	Total	2425 (100.0)	26.5 (8.0-134.8)	0.0-14477.6	1185 (48.9)	303 (12.5)	240 (9.9)	318 (13.1)	108 (4.5)	271 (11.2)

^a: positional error was determined by calculating Euclidian distance (in meter) between addresses located by each automatic method (method A

545 and method B) of geocoding regarding to a reference method (method R).

^b: the administrative division of territories in France did not allow obtaining addresses geocoded to district and state levels (level 2), county (level 3) and intersection of streets (level 7). No addresses have been geocoded to the country level.

Table 2: Agreement in accuracy level (Cohen's Kappa coefficient) for methods A and B (automatic geocoding) in comparison with method R (manual reference method).

Method R	Method A						Method B					
	City (N addresses)	Street segment (N addresses)	Address (N addresses)	Total (N addresses)	Cohen's kappa coefficient	Postal code (N addresses)	Street segment (N addresses)	Address (N addresses)	Total (N addresses)	Cohen's kappa coefficient		
Overall	Town Hall	147	20	14	181		133	49	5	187		
	Street segment	212	309	140	661	0.60	119	439	131	689	0.61	
	Address	46	34	1294	1374		125	70	1354	1549		
	Total	405	363	1448	2216		377	558	1490	2425		
For urban addresses	Town Hall	62	10	11	83		54	27	5	86		
	Street segment	94	166	130	390	0.56	64	226	126	416	0.52	
	Address	44	26	1265	1335		114	63	1331	1508		
	Total	200	202	1406	1808		232	316	1462	2010		
For rural addresses	Town Hall	85	10	3	98		79	22	0	101		
	Street segment	118	143	10	271	0.39	55	213	5	273	0.54	
	Address	2	8	29	39		11	7	23	41		
	Total	205	161	42	408		145	242	28	415		
For 1990-2000 period addresses	Town Hall	133	17	13	163		125	45	5	175		
	Street segment	182	278	120	580	0.61	112	396	117	625	0.60	
	Address	45	29	1141	1215		121	67	1216	1404		
	Total	360	324	1274	1958		358	508	1338	2204		
For 2000-	Town Hall	14	3	1	18	0.56	8	4	0	12	0.70	

2008 period addresses	Street segment	30	31	20	81		7	43	14	64	
	Address	1	5	153	159		4	3	138	145	
	Total	45	39	174	258		19	50	152	221	

550 ^a. Assignment of urban and rural status of addresses was based on definitions established by the French national institute for statistics and economic studies (INSEE)

5. Discussion

In the GEO3N project, airborne dioxin exposure was estimated using a GIS-based metric that required the retrospective inventory of industrial sources, the estimation of their dioxin emission intensity, the geocoding of the participants' residential history and of the industrial sources between 1990 and 2008 and the computation of environmental and meteorological parameters in the GIS. Parameters included in the GIS-based exposure metric were determined and combined based on a literature review and on a comparison with a dioxin atmospheric dispersion modeling. Several parameters and combinations were tested and the final metric included parameters' combination that showed the best agreement with dispersion modeling. The list of parameters included in the GIS-based metric might not be exhaustive, and we did not consider wind speed, exhaust smoke velocity and stack height which did not further improve our metric and were not available for every industrial source. However, weighted kappa coefficients showed "substantial" to "almost perfect" agreement with dispersion modeling.

In early studies, GIS have been used to compute exposure surrogates based on distance between study population and contaminant sources. The use of GIS has recently increased and advanced methods have been developed, integrating meteorological and topographical data, residential history as well as characteristics of industrial sources (Hoek et al. 2001; Pronk et al. 2013; Vienneau et al. 2009; White et al. 2009; Yu 2006). GIS represents now a potential method to minimize exposure misclassification in air pollution exposure assessment (Zou et al. 2009). We developed a GIS-based metric in this way, filling methodological gaps of the existing literature to improve accuracy of airborne dioxin exposure estimates and assess more reliable associations with breast cancer.

An exhaustive and retrospective inventory of industrial sources likely to emit or to have emitted dioxins between 1990 and 2008 in the Rhône-Alpes region was carried out and led to the identification and characterization of 302 industrial sources from many sectors: waste incineration, metal production, heat and power generation, production of mineral, chemical and consumer goods and crematoria.

In the absence of national measured dioxin concentrations covering the entire study period, and given the exhaustive and retrospective nature of the exposure assessment, annual dioxin emission intensities of the industrial sources were estimated using the UNEP Toolkit, a standardized tool that considers industrial sectors and technical characteristics to assign an emission factor to each source inventoried (UNEP 2013). This method did not allow an accurate assessment of retrospective emissions as continuous data measurement would do; however, it allowed the estimation of magnitudes of dioxin emissions for every source with a homogeneous and reproducible methodology, leading to the relative ranking of sources in order to estimate the relative exposure of each of the study participants.

Other punctual and non-industrial sources can emit relatively high amounts of dioxins, locally and on a short time scale. These concern biomass fires and volcanic eruptions, biomass and manufactured good burnings, cable burning, outdoor burning and illegal landfills in the early 1990s. These sources could not be considered in the estimation of the airborne dioxin exposure due to the difficulty of their retrospective inventory, their geolocalization and the estimation of their dioxin emissions. Dioxin emissions from road traffic were not considered in the estimation of airborne dioxin exposure because their contribution was negligible and remained stable over the study period. Domestic activities can generate dioxins, such as heating, chimney fire, cooking methods (stove, wood stove, and barbecue) or personal burning of green waste. E3N follow-up questionnaires did not provide the information. However, these items will be part of a new self-administered questionnaire sent to the E3N women

and retracing their lifetime residential history, in order to consider domestic activities in the assessment of environmental exposure within the E3N cohort. This aspect will be developed in Chapter IV. Although they poorly contributed to airborne dioxin exposure compared to industrial sources, this could have led to underestimation of the exposure for the participants exposed to these extra-sources and to misclassification bias.

Residential history of the participants was geocoded with an automatic geocoding method (Article #2). The addresses were not recorded to be geocoded and used in the estimation of environmental exposure. However, their accuracy was evaluated and addresses were considered precise enough to limit misclassification bias (Han et al. 2013).

Strengths in the airborne dioxin exposure assessment included the use of a GIS and its application on a large area, over a long and retrospective period, at the individual subject's address and considering the residential history over the study period (Brody et al. 2007; Cordioli et al. 2013). Exposure assessment was based on a multi-sources approach, considering waste incineration, metal and cement industries etc. and the evolution over time of the facilities' technical characteristics. In the absence of dioxin monitoring data, emission intensity of the industrial sources was estimated using a standardized tool. Residential addresses and industrial source locations were geocoded with an automatic geocoding method whose accuracy was evaluated (Han et al. 2013). The GIS-based metric was validated through comparison with dispersion modeling. Finally, this GIS-based metric will be applied in future epidemiological research to investigate associations between environmental exposures and disease's risk.

In Chapter III, we developed and validated a GIS-based metric to estimate airborne dioxin exposure of the women from the E3N cohort; the next objectives are to apply the GIS-based metric on the E3N women living in the Rhône-Alpes region and to estimate the risk of breast cancer associated with this exposure.

CHAPTER IV. GIS-BASED AIRBORNE DIOXIN EXPOSURE AND BREAST CANCER RISK

Assessment of airborne dioxin exposure and estimation of breast cancer risk

1. Introduction

Dioxins are emitted from industrial activities such as metal and cement industry and incineration, and released into the environment. Evidence regarding dioxin exposure and breast cancer risk mostly involved populations exposed accidentally or occupationally to high levels of dioxins (Bertazzi et al. 1993; Brody et al. 2007; Manuwald et al. 2012; Pesatori et al. 2009; Warner et al. 2002, 2011). Other studies investigated the association among residents living in the vicinity of chemical plants or waste incinerators (Fabre et al. 2008; Revich et al. 2001; Viel et al. 2008). Results of published studies remained inconsistent, due to study limitations such as poor exposure assessment and no adjustment for breast cancer risk factors (Cordioli et al. 2013).

Dioxins are non-genotoxic and tumor-promoter (Lamartiniere 2002; Teitelbaum et al. 2015) so the impact of time-varying and late exposures, as well as specific windows of breast susceptibility, need to be considered in the investigation of retrospective environmental exposure and breast cancer risk (Vacek 1997).

Objectives and methods

The objectives of this chapter were to assess airborne dioxin exposure and to estimate the risk of breast cancer associated with this exposure in a case-control study nested within the E3N cohort (GEO3N study population). The methodology for exposure assessment and risk analysis was first elaborated in the Rhône-Alpes region, France. Therefore, this work was restricted to 429 breast cancer cases and 786 matched controls that had permanently lived in this region over the study period (1990-2008).

Retrospective airborne dioxin exposure was estimated, at the individual level, by applying the GIS-based metric previously developed in Chapter III. At the time of analysis, we could consider all industrial sources inventoried in the Rhône-Alpes region between 1990 and 2008 except for hospital waste incinerators for which no data were available on dioxin emission intensity and location at the time of analysis. Residential history and industrial sources were geocoded in the GIS to compute residence-to-source distance and prevailing wind frequency at each location. The UNEP toolkit was used to estimate emission intensities of the industrial sources (UNEP 2013).

The association between airborne dioxin exposure and breast cancer risk was first estimated using conditional logistic regression models. Breast cancer risk estimates were reported for an increase of 1 standard deviation and by quintiles of cumulative airborne dioxin exposure. Models were adjusted for breast cancer risk factors and potential confounders; models were also further adjusted for estimated dietary dioxin exposure (Article #1). Subgroup analyses were performed according to menopausal status and hormone-receptor status of the tumors. Further analyses were conducted among subgroups defined by BMI, breastfeeding and maternal age, using unconditional logistic regression to retain all study subjects; models were then further adjusted for matching factors. We also estimated time-dependent odds ratios according to time before diagnosis by modeling weight functions, as described in Hauptmann et al. 2000; this work was achieved in collaboration with Karen Leffondré and Emilie Lévêque from the team “Epidemiology and Biostatistics” (Inserm U1219, ISPED, Université de Bordeaux) and was supported by a mobility grant from the CLARA.

2. Article #3

Airborne Dioxin Exposure and Breast Cancer Risk in a case-control study nested within the French E3N Prospective Cohort: a GIS-based approach

Aurélie MN Danjou^{1,2}, Thomas Coudon^{1,2}, Karen Leffondré³, Emilie Lévêque³, Elodie Faure¹, Delphine Praud¹, Marie-Christine Boutron-Ruault^{4,5,6}, Thierry Philip^{1,2}, Laure Dossus⁷, Françoise Clavel-Chapelon^{4,5,6} and Béatrice Fervers^{1,2}

¹ Département Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France

² Université de Lyon, Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69100 Villeurbanne, France

³ Université de Bordeaux, Institut de Santé Publique, d'Épidémiologie et de Développement, Centre Inserm U1219 Epidemiology and Biostatistics, 146 rue Léo Saignat, 33076 Bordeaux, France

⁴ Inserm, Center for Research in Epidemiology and Population Health (CESP), U1018, team "Lifestyle, genes and health", 114 rue Edouard-Vaillant, 94805 Villejuif Cedex, France

⁵ Université Paris-Sud, UMRS 1018, Villejuif, France

⁶ Gustave Roussy, F-94805 Villejuif, France

⁷ Nutritional Epidemiology Group, Nutrition and Metabolism Section, International Agency for Research on Cancer, 150 cours Albert Thomas, 69372 Lyon Cedex 08, France

Reference:

Danjou AMN, Coudon T, Leffondré K, Lévêque E, Faure E, Boutron-Ruault MC, Philip T, Clavel-Chapelon F, Dossus L, Fervers B 'Airborne dioxin exposure and breast cancer risk in a case-control study nested within the E3N prospective cohort: a GIS-based approach. *To be submitted to Environmental Health.*

Airborne Dioxin Exposure and Breast Cancer Risk in a case-control study nested within the French E3N Prospective Cohort: a GIS-based approach

Aurélie M N Danjou ^{1,2}, Thomas Coudon ^{1,2}, Karen Leffondré ³, Emilie Lévêque ³, Elodie Faure ¹, Delphine Praud ¹, Marie-Christine Boutron-Ruault ^{4,5,6}, Thierry Philip ^{1,2}, Laure Dossus ⁷, Françoise Clavel-Chapelon ^{4,5,6} and Béatrice Fervers ^{1,2*}

Authors' affiliations

¹ Département Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France

² Université de Lyon, Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69100 Villeurbanne, France

³ Université de Bordeaux, Institut de Santé Publique, d'Épidémiologie et de Développement, Centre Inserm U1219 Epidemiology and Biostatistics, 146 rue Léo Saignat, 33076 Bordeaux, France

⁴ Inserm, Center for Research in Epidemiology and Population Health (CESP), U1018, team «Lifestyle, genes and health», 114 rue Edouard-Vaillant, 94805 Villejuif Cedex, France

⁵ Université Paris-Sud, UMRS 1018, Villejuif, France

⁶ Gustave Roussy, F-94805, Villejuif, France

⁷ Nutrition and Metabolism Section, International Agency for Research on Cancer, 150 cours Albert Thomas, 69372 Lyon Cedex 08, France

* Correspondence to Béatrice Fervers, Département Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France. Telephone: +33(0)478782801. E-mail: beatrice.fervers@lyon.unicancer.fr

Manuscript in preparation, to be submitted to Environmental Health

Abstract

Background:

Dioxins are environmental by-products emitted by industrial chlorinated combustion processes. Dioxins act as endocrine disruptors and are suspected to be involved in breast cancer (BC) etiology. 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) was classified as a Group 1 carcinogen by the International Agency for Research on Cancer. However, results on the association with BC remain inconsistent. TCDD is considered as non-genotoxic carcinogen and acts as tumor promoter. The timing of exposure may affect findings in epidemiological studies. We aimed at investigating BC risk associated with airborne dioxin exposure, by developing a geographic information system (GIS)-based exposure metric and considering temporal dimensions of exposure in the risk estimate.

Methods:

We designed a case-control study nested within the French E3N prospective cohort. Between 1990 and 2008, we identified 429 invasive BC cases that were matched to 786 randomly selected controls. The study population permanently lived in the Rhône-Alpes region, France, over the study period. Assessment of dioxin exposure was based on a detailed inventory of dioxin emitting sources and residential history of the study subjects. Exposure was evaluated at the individual address level with a GIS-based exposure metric including proximity to and technical characteristics of dioxin emitting sources, exposure duration and wind direction. We first studied BC risk in relation to cumulative exposure using conditional logistic regression models. We then modeled a time-dependent weight function as a cubic B-spline and estimated time-dependent odds ratios (OR) for BC.

Results:

A one standard deviation (SD) increase in airborne dioxin exposure was not associated with overall breast cancer (OR=0.99, 95%CI: 0.87-1.13). Statistically significant inverse associations were observed among women with a BMI \leq 21.8 kg/m², having had their first child before 24 years of age and with a history of breastfeeding. OR for breast cancer according to time before diagnosis were similar at each time-point, and corresponded to the OR from the standard model with cumulative airborne dioxin exposure.

Conclusions:

No association was observed between airborne dioxin exposure and BC risk overall. The results suggest a decreased BC risk associated with low BMI, breastfeeding and young age at first child. Confirmation of our findings is required in a larger population. Our study may

Author's draft

65 provide a new tool for the assessment of long-term exposure with the GIS-based metric as
well as new ways of considering temporal dimensions for environmental exposures in disease
risk assessment.

Keywords

70 Breast cancer, endocrine disruptors, dioxins, environmental exposure, geographic information
system, timing of exposure, weight function, tumor receptor

Background

75 Dioxins include polychlorinated dibenzo-*para*-dioxins (PCDD) and polychlorinated
dibenzofurans (PCDF) and belong to the halogenated polycyclic aromatic hydrocarbons
(PAHs). Dioxins are environmental by-products formed from combustion processes involving
a chlorine source. The main sources of dioxin release into the environment include chemical
manufacturing of chlorinated products and industrial activities from metallurgy, steel and
80 municipal solid waste incineration (MSWI) [1–4]. Dioxin emissions have been greatly
reduced since the 1990s, because of changes in regulation and technological processes [4,5].
In France, emissions have been reduced by a factor 1000 between 1990 and 2008 [6]. Dioxin
exposure in the general population mainly occurs from food consumption, but also by means
of inhalation. Persistent in the environment and the human body, they accumulate in adipose
85 tissues [7].

 2,3,7,8-tetrachlorodibenzo-*para*-dioxin (TCDD) is the most toxic dioxin congener and
has been classified as carcinogenic to humans (group 1) by the International Agency for
Research on Cancer (IARC), with sufficient evidence for all cancers combined [8]. In
addition, TCDD acts as an endocrine disruptor *via* binding to the aryl hydrocarbon receptor
90 (AhR) [9,10], and may contribute to the development of hormone-dependent malignancies
such as breast cancer. Experimental research also demonstrated that dioxins are non-
genotoxic carcinogens with tumor-promoting properties and act through activation of cellular
replication, apoptosis and increase of oxidative stress leading to DNA damage and mutations
[7,11].

95 Evidence regarding dioxin exposure and breast cancer risk mainly comes from
populations accidentally and occupationally exposed, and from populations living in the
vicinity of industrial facilities, and remains inconsistent. Following the 1976 Seveso (Italy)
industrial accident, results from the cohort of local residents in the contaminated zones
showed no association with breast cancer risk [12,13] and no association was observed
100 between TCDD serum levels measured in blood samples provided after the accident and
breast cancer risk among women who were 0- 40 years old at the time of the accident and
lived in the most contaminated zones [14]. Epidemiological studies conducted among cohorts
of herbicide and insecticide workers showed higher mortality and incidence of female breast
cancer [15,16]. An elevated breast cancer mortality was reported among Russian women
105 living near a chemical plant, compared to women from a non-contaminated area [17]. In
France, a statistically significant increase in breast cancer risk was observed among women

Author's draft

from an ecological study of residents exposed to the emissions from MSWIs [18]. In a French case-control study conducted around a MSWI, a statistically significant decrease in breast cancer risk was found among women over 60 years of age and living in the most exposed zone [19]. This inconsistency in published results may be explained by differences in methodology (study design and exposure assessment) as well as some specific study limitations such as lack of adjustment for confounding by established risk factors, inadequate exposure assessment and small numbers of cases [20].

The long-term nature of the airborne dioxin exposure may imply variations in exposure intensities over time, reinforcing the need to consider specific windows of exposure [21]. Moreover, women may experience multiple time-variable windows of breast susceptibility (prenatal period, puberty, pregnancy and menopause) in which dioxins could affect hormonal pathways and induce the development of breast cancer later in life [22,23]. Given their tumor-promoter properties, studying the association between airborne dioxin exposure and breast cancer risk should consider these specific windows of susceptibility and the impact of time-varying and late exposures.

Recently, we examined the relationship between estimated dietary dioxin exposure and breast cancer risk in the French E3N (Étude Épidémiologique auprès des femmes de la Mutuelle Générale de l'Éducation Nationale) prospective cohort. No association was observed overall, except for a statistically significant decrease in postmenopausal hormone-independent breast cancer risk among women most exposed [24]. Here, we investigated breast cancer risk associated with airborne dioxin exposure, in a case-control study nested within the E3N cohort, improving the methodology for the assessment of low-dose airborne dioxin exposure. Furthermore, we applied statistical methods allowing to consider temporal dimensions of exposure in the risk estimate.

Methods

The E3N cohort study

135 The E3N study is an ongoing prospective cohort, initiated to identify female cancer risk
factors, such as dietary, hormonal and reproductive factors [25]. The cohort includes 98,995
French female volunteers, born between 1925 and 1950 and national health insurance
subscribers, most of whom are teachers. The study protocol was approved by the French
National Commission for Data Protection and Privacy. The study was initiated in 1989-1991
140 when participants returned the first self-administered questionnaire and provided written
informed consent for data collection. Since then, participants completed self-administered
questionnaires, mailed every 2 to 3 years, on health status, medical history and a large number
of lifestyle factors. Epidemiological data were completed with 25,000 blood samples and
47,000 saliva samples. Participants' home addresses were also recorded in the first and fifth to
145 ninth questionnaires; their zip code was recorded in the third and fourth questionnaires; there
was no address kept in the database for the second questionnaire. Participants' place of birth
(zip code and commune) was obtained from the first questionnaire and assigned an
urban/rural status; definition of the urban/rural status has been previously described in [26].
The E3N cohort is the French component of the European Prospective Investigation into
150 Cancer and nutrition (EPIC) study [27,28].

Study population

For this nested case-control study, we selected participants who filled in their home address at
baseline, who lived in the Rhône-Alpes region, France between 1990 and 2008 and who had
155 not reported cancer at baseline. Data on health status and medical history of participants,
including cancer occurrence, were collected and updated throughout the follow-up. Between
1990 and 2008, 429 primary invasive breast cancer cases were reported, of which 93% have
been validated with pathology reports; cases not confirmed were considered as breast cancer
cases, the proportion of false-positive self-reported breast cancer cases being less than 5% in
160 the E3N cohort. Information on estrogen receptor (ER) and progesterone receptor (PR) status
was obtained from the pathology reports: n=244 tumors were ER positive (ER+); n=80 were
ER-negative (ER-); n=215 were PR positive (PR+); and n=108 were PR-negative (PR-).

According to an incidence density sampling plan, up to two controls per case were
randomly selected and matched for age (± 1 year), department of residence, date (± 3 months)
165 and menopausal status at blood collection for the participants who donated blood, or at

baseline for the participants without a blood sample who were then matched for existence or not of a saliva sample. A total of 786 controls were selected.

Assessment of airborne dioxin exposure

170 *Inventory and characterization of dioxin emitting sources*

A complete retrospective inventory of industrial sources likely to emit or to have emitted dioxins between 1990 and 2008 in the Rhône-Alpes region was performed. The industrial sources were identified through national databases, institutional information sources and structured questionnaires, and information on technical and process characteristics (e.g. stack
175 height, exhaust flow rate and flue gas cleaning technologies), operation periods and rates, input materials and geographic location of the chimney stack were collected.

A total of 224 dioxin emitting sources operated in the Rhône-Alpes region from 1990 to 2008, corresponding to 290 distinct operation periods according to the evolution of technical characteristics. The predominant sector was waste incineration (N=122 distinct
180 operation periods-sources), followed by production of mineral (N=88), heat and power generation (N=40), metal production (N=26), crematoria (N=13) and chemicals and consumer goods (N=1).

Using the 2013 Standardized Toolkit for Identification and Quantification of Dioxin and Furan Releases of the United Nation Environment Program [29], the industrial sources
185 were classified according to their main activity sector and their technical characteristics to determine a dioxin emission factor (g-TEQ/t). For each industrial source, the annual dioxin emission intensity (g-TEQ/year) was estimated by multiplying the emission factor by the operation rate.

190 *Geocoding of participants' and industrial sources' addresses*

From 1990 to 2008, the residential history of the participants was extracted from the E3N follow-up questionnaires and geolocated using the ArcGIS Software (ArcGIS Locator version 10.0, Environmental System Research Institute – ESRI, Redlands, CA, USA) and the address
195 database from the National Geographic Institute (IGN) (Faure *et al.* submitted). Geocoding was performed by a trained technician blinded to the case-control status of the participants. In addition, each industrial source inventoried was manually geolocalized in the GIS at the location of the flue-gas stack.

Exposure assessment

200 Airborne dioxin exposure was estimated for each participant at the individual address level
 using a GIS-based exposure metric. Relevant parameters for inclusion in the GIS-based metric
 were derived from the literature [30–35] and comparison with dioxin atmospheric dispersion
 modeling using SIRANE, an urban Gaussian dispersion model [36,37], for three distinct years
 (1996, 2002 and 2008) in two areas with distinct meteorological data and varying emission
 205 intensities, including the following parameters: the Toolkit-based annual dioxin emission
 intensity of the industrial sources; an impact zone of dioxin emissions around each industrial
 source, named “buffer” (3; 5 and 10 km); a distance decline pattern ($1/d$; $1/d^{1.5}$; $1/d^2$; $1/d^3$; e^{-d} ;
 e^{-d}/d and $e^{-d} \cdot d$); and prevailing wind frequency. For the latter, hourly data on prevailing wind
 direction were provided by *Météo France* ® and averaged over the year for each year between
 210 1990 and 2008. Each industrial source was assigned to a meteorological station based on local
 meteorology and topography. Buffers around each industrial source were divided into equal
 wind direction segments, which were assigned an annual wind direction frequency. Distinct
 segment sizes were tested: 90°; 30° and 10°.

Agreement between classifications into quintiles of the dispersion modeling and the
 215 GIS-based metric was calculated with weighted kappa coefficients and their 95% confidence
 interval (CI). Kappa statistics between 0.61 and 0.80 represent “substantial agreement” and
 kappa statistics greater than 0.81 represent “almost perfect agreement” [38]. Weighted kappa
 coefficients ranged from 0.71 (0.67-0.76) to 0.84 (0.79-0.88) and were consistent across
 periods and areas (rural/urban) for an impact zone of 10 km, a residence-to-source distance
 220 decline of $1/d^2$ and a prevailing wind frequency divided by 10° segments with weighted
 contribution of adjacent segments. Additional inclusion of exhaust smoke velocity, stack
 height and wind speed did not further improve the weighted kappa coefficients. The selected
 parameters were integrated in the final GIS-based exposure metric using the following
 formula:

$$225 \quad \text{GIS - based metric} = \sum_{j=1}^J \sum_{i=1}^I \left(EI \times t \times \frac{1}{d^2} \times F \right)$$

Where j is the place of residence ($j=1, \dots, J$), i is the industrial source ($i=1, \dots, I$), EI is the
 annual dioxin emission intensity (in g-TEQ/year), t is the exposure duration (in year), d is the
 residence-to-source distance (in m) and F is the prevailing wind frequency. Airborne dioxin
 exposure was expressed in $\mu\text{g-TEQ}/\text{m}^2$ and computed for each calendar year from 1990 to
 230 2008.

Statistical analyses

Baseline characteristics of the participants were described according to their case-control status using mean and standard deviation (SD) for quantitative variables and frequency and percentage for qualitative variables. Cases and controls' characteristics were compared with conditional logistic regression.

Airborne dioxin exposure was summarized into a cumulative exposure metric from baseline to the index date (date of diagnosis of the case in the case-control(s)' pair) by adding the calendar values. Airborne dioxin exposure was also categorized *a priori* according to quintiles of its distribution among controls.

Odds ratios (OR) and 95% CI for invasive breast cancer were estimated for an increase of 0.88 $\mu\text{g-TEQ}/\text{m}^2$ (that is 1 SD in the study population) and for quintiles of the cumulative exposure metric with the first quintile as the referent group, using conditional logistic regression models. Tests for linear trend across quintiles were performed and derived from the Wald test of the models with the continuous variable. All models were adjusted for matching factors and established breast cancer risk factors and potential confounders: education (undergraduate/ postgraduate with a 1- to 4-year university degree/ postgraduate with a 5+ year university degree); height (cm); body mass index (BMI, cut off: 25, 30 kg/m^2); smoking status (never smoked/ ex-smoker/ current smoker); alcohol intake (cut off: 0, 6.1 g/day); physical activity (cut off: 25.3, 37.3, 57.6 metabolic equivalent task-hour per week (MET-h/w)); previous family history of breast cancer (yes/no); previous history of personal benign breast disease (yes/no); previous mammography (yes/no); age at menarche (cut off: 12, 14 years); use of oral contraceptives (ever/never); use of progestin before menopause (ever/never); age at first full-term pregnancy combined with parity (≤ 2 children and age at first birth < 30 years/ ≤ 2 children and age at first birth ≥ 30 years/ ≥ 3 children); breastfeeding (ever/never); use of menopausal hormone therapy (MHT, yes/no, among post-menopausal women only); age at menopause (cut off: 47, 54 years, among post-menopausal women only); rural/urban status of the subjects' birthplace (rural/urban). For adjustment variables with less than 5% missing data, the latter were replaced by their modal or median value in the control population. Alcohol intake, age at first full-term pregnancy combined with parity and status of birthplace had more than 5% of missing data, thus a category "missing data" was generated. We estimated invasive breast cancer risk according to hormone-receptor status (ER and PR) of the tumors, excluding cases for whom the status was missing (N=105 for ER; N=106 for PR) from the corresponding analyses. We also studied the association among subgroups defined by menopausal status (premenopause; postmenopause). In addition, models were

Author's draft

further adjusted for estimated dietary dioxin exposure, which was assessed in the E3N cohort in a previous study [24].

Several interactions were tested: between BMI and menopausal status, because of opposite associations between BMI and risk of breast cancer in pre- and post-menopause [39];
270 between BMI (cut off: 21.8 kg/m², median value among controls at baseline) and cumulative airborne dioxin exposure, because of the storage of dioxins in adipose tissue; between breastfeeding (ever/never) and cumulative airborne dioxin exposure, because of previously observed inverse association with breastfeeding [40,41]; between age at first full-term pregnancy (cut off: 24 years; median value among controls) and cumulative airborne dioxin
275 exposure, because of previously observed positive association with age and maternal age due to dioxin's persistence and accumulation in the human body [41,42] and potential inverse association with parity [42,43]. All tests for interaction were derived from the likelihood ratio test comparing the models with and without interaction term. Association between invasive breast cancer risk and cumulative airborne dioxin exposure was examined in the above-
280 defined subgroups, using unconditional logistic regression and adjusting for matching factors, in order to retain all subjects in the models.

Furthermore, we investigated the time-dependent impact of airborne dioxin exposure on breast cancer risk, based on an approach first developed by Hauptmann *et al.* [44]. A time-dependent weight function $w(t)$ was estimated according to time t prior to breast cancer
285 diagnosis and assuming equal 1-year time intervals. The weight function assigned weights to past exposures and was modeled as a cubic B-spline on $[0, T]$, where T was the maximum exposure time ($T=19$). We estimated time-dependent OR for breast cancer within unconditional logistic regression models that included the time-weighted airborne dioxin exposure (for an increase of 0.1 $\mu\text{g-TEQ/m}^2$), adjusting for matching factors and status of the
290 birthplace. Because B-splines are conditional on inner knots, the number of knots was determined in order to minimize the Akaike information criteria (AIC). Nonparametric bootstrap sampling was performed to obtain 95% confidence interval for the estimated weight function.

All P -values were two-sided and the significance level was set at 0.05. We used the
295 SAS statistical software version 9.4 (SAS Institute Inc., Cary, NC, USA) and the R software for data analysis.

Results

300 Baseline characteristics of the study population are shown in Table 1. Alcohol consumption, age at first full-term pregnancy over 30 years old, family history of breast cancer and personal history of benign breast disease were more frequent among cases than controls. There was no difference in dietary dioxin exposure between breast cancer cases and controls ($p=0.25$).

Table 1 on baseline characteristics should appear here.

305 The average cumulative airborne dioxin exposure was estimated at $0.14 \pm 1.2 \mu\text{g-TEQ/m}^2$ (range: 0 to $24.4 \mu\text{g-TEQ/m}^2$) for breast cancer cases and $0.13 \pm 0.6 \mu\text{g-TEQ/m}^2$ (range: 0 to $24.4 \mu\text{g-TEQ/m}^2$) for controls. The average exposure duration between recruitment and diagnosis among cases was 7.2 ± 4.6 years (range: 1 month to 18 years). No association was observed between cumulative airborne dioxin exposure and breast cancer risk overall in conditional logistic regression analyses after adjustment for individual breast cancer risk factors (Table 2). Odds ratio for breast cancer was 0.99 (95%CI: 0.87-1.13) for an increase of $0.88 \mu\text{g-TEQ/m}^2$ in exposure. There was no statistically significant linear trend between quintiles of cumulative airborne dioxin exposure and breast cancer risk (P for trend=0.18). There was no association between cumulative airborne dioxin exposure and hormone-receptor defined breast cancer risk. However, a decrease in ER-negative breast cancer, although statistically non-significant, was consistently observed across quintiles of cumulative airborne dioxin exposure (OR from 0.56, 95%CI: 0.24-1.31 for Q1 vs Q2 to 0.70, 95%CI: 0.29-1.69 for Q1 vs Q5; P for trend=0.14). When stratifying by menopausal status, no statistically significant association was observed between cumulative airborne dioxin exposure and breast cancer risk. Further adjustment on estimated dietary dioxin exposure did not change the results (data not shown).

Table 2 should appear here.

325 There was no effect modification by BMI, age at menarche, parity, breastfeeding and age at first full-term pregnancy. However, statistically significant inverse associations between cumulative airborne dioxin exposure and breast cancer risk were found among women with a BMI $\leq 21.8 \text{ kg/m}^2$ (OR for Q1 vs Q5: 0.49, 95%CI: 0.24-0.98; P for trend=0.01), women having had their first child before 24 years of age (OR for Q1 vs Q5: 0.15, 95%CI: 0.06-0.37; P for trend=0.0003) and women with a history of breastfeeding (OR for Q1 vs Q5: 0.41, 95%CI: 0.22-0.79; P for trend=0.01) (Table 3).

330 *Table 3 should appear here.*

We specified models with one knot. For an increase of $0.1 \mu\text{g-TEQ/m}^2$ of cumulative

Author's draft

airborne dioxin exposure, odds ratios for breast cancer according to time before diagnosis were 1.00 with similar 95% CI at each time-point. Although the time-dependent risk estimates were statistically non-significant, the lower bound of the 95% CI suggested a slightly inverse association. The odds ratios were similar to the corresponding odds ratios for overall breast cancer risk from the standard model with cumulative airborne dioxin exposure (Table 4).
335 *Table 4 should appear here.*

Discussion

340

Among women from the E3N cohort, we observed overall no association nor time-dependent effects of cumulative airborne dioxin exposure and breast cancer risk. Suggestive decreased breast cancer risk estimates were observed among lean women, women with a younger age at first full-term pregnancy and women with a history of breastfeeding.

345

Epidemiological studies published so far on the relationship between dioxin exposure and breast cancer risk yielded inconsistent results. Contrary to our results, weak increases in breast cancer risk [13,45] and mortality rate [15,16] have been observed in some studies. However, differences may be explained by the level of dioxin exposure which was higher among accidentally and occupationally exposed populations than in our study population.

350

Evidence from studies conducted in the general population has been contradictory and focused on women living in the vicinity of industrial facilities [17–19]. Elevated mortality rates for breast cancer were found among women living near a chemical plant in a Russian region, compared to a non-contaminated region [17]; again, dioxin exposure levels were higher than in our study population. A French ecological study showed a positive association

355

between living under the plume of a MSWI and breast cancer occurrence [18]. Although the study was based on a large number of women, there was no adjustment for individual breast cancer risk factors, limiting interpretation of results. Also, a French population-based case-control study, conducted in the vicinity of a MSWI, reported a decreased breast cancer risk among women over 60 years of age and living in the most dioxin exposed zone compared to

360

women from the lowest exposed zone [19]. However, analyses were not adjusted for breast cancer risk factors. Most of the above studies did not assess dioxin exposure at the individual address level [13,17–19] and none considered dietary dioxin exposure in the statistical analysis. Furthermore, our findings are in agreement with the results from a previous study conducted in the E3N cohort, in which we observed no association between estimated dietary dioxin exposure and overall breast cancer risk [24].

365

Although they were not statistically significant, our results suggested that low-dose cumulative airborne dioxin exposure may be associated with decreased ER-negative breast cancer risk. This observation is consistent with the statistically significant decrease in postmenopausal ER-negative PR-negative breast cancer risk found among women from the E3N cohort with the highest estimated dietary dioxin exposure [24]. Moreover, previous studies reported inverse associations between endocrine disrupting chemicals and ER-negative breast cancer risk. In a Danish nested case-control study, higher adipose tissue

370

Author's draft

concentrations of organochlorines were associated with a statistically significant lower risk of postmenopausal ER-negative breast cancer [46]. In a population-based case-control study in Long Island, a statistically significant lower risk of ER-negative PR-negative breast cancer was observed among women with higher serum levels of polychlorinated biphenyls (PCB) [47]. Additional studies reported inverse associations between organochlorine pesticide and PCB exposures and overall breast cancer risk; however, they did not reach statistical significance [48–50]. Biological mechanisms have been suggested to explain this inverse association. *In vitro* studies demonstrated the antiproliferative action of TCDD in ER-negative breast cancer cells, through AhR-dependent [51,52] and AhR-independent pathways [53]. Moreover, animal studies described a potential role of TCDD in the inhibition of mammary tumor formation [54].

Statistically significant decreased breast cancer risks were observed among lean women, women with a younger maternal age and women with a history of breastfeeding – although no effect modification was found. Dioxin serum concentrations have been shown to decrease with longer breastfeeding and parity [41–43], and increase with maternal age [42] and BMI [55], and may explain the observed inverse associations. However, the effects of pregnancy and lactation on the elimination of persistent organic pollutants are complex and remain insufficiently understood [56]. Correlations between BMI and dioxin concentrations were inconsistent across studies [57,58] and no increase in breast cancer risk was observed among overweight women in our study. Moreover, our subgroups analyses were based on small numbers of participants, possibly yielding unstable risk estimates. We cannot exclude that some of the statistically significant findings may have occurred by chance.

Time-dependent risk estimates showed no association between cumulative airborne dioxin exposure and overall breast cancer risk, and the risk estimates were similar at each time-point. The estimated weight function was unstable and imprecise, mostly because this statistical method is sensitive both to the number of participants, to the number of knots and to the exposure values (which ranged from pg to $\mu\text{g-TEQ/m}^2$ in our study population). Thus our results require further research and confirmation in a larger study population.

Our study has some limitations. The dioxin exposure was not directly measured in blood samples or adipose tissue. Although the E3N women provided blood samples in 1995–1997, the volume in storage was insufficient to measure dioxin levels, and a new sampling would not be relevant regarding the etiology of breast cancer. Indirect measures include the use of land use regression, dispersion modeling and GIS-based methods, and their use in epidemiological studies have grown in the last years [20]. Because of the lack of airborne

dioxin monitoring data in space and time, the sharp decrease in dioxin emissions in France over the study period, as well as the number of dioxin emitting sources included in our study, we could not apply land use regression or dioxin dispersion models. Therefore, airborne
410 dioxin exposure was assessed through a GIS-based dioxin exposure metric that allowed the estimation of long-term exposure at the individual address level. The GIS included parameters known to influence individual airborne dioxin exposure according to the literature: residence-to-source proximity, dioxin emission intensity of industrial sources, exposure duration and wind direction. Exposure classification using the GIS-based metric showed “substantial” to
415 “almost perfect” agreement with the SIRANE dispersion modeling across different settings. We were able to consider the residential history of the participants over the study period and we implemented a standardized method for the estimation of dioxin emissions using the UNEP Toolkit [29]. Geocoding of the participants’ residential addresses and industrial facilities was achieved through an automatic method whose accuracy was assessed in a
420 previous study (Faure *et al.* submitted).

We estimated airborne dioxin exposure considering major industrial dioxin emitting sources, but not minor sources or traffic-related exposure. However, to our knowledge, this was the first study to assess airborne dioxin exposure based on a multi-source approach, considering waste incineration, metal and cement industries, etc. and the evolution over time
425 of the facilities’ technical characteristics. The contribution of traffic to airborne dioxin exposure was estimated to be negligible (<2%) [6] and remained stable over the study period.

Emissions from domestic activities, e.g. heating, chimney fire, cooking methods (e.g. stove, wood stove and barbecue) as well as backyard burning of domestic and green wastes, can contribute to airborne dioxin exposure and have been associated with risk of breast cancer
430 [59]. Also, the lack of past residential history and historical dioxin exposure estimates (before 1990) may have resulted in exposure misclassification likely to contribute to imprecise risk estimates drawn toward unity [60]. Information on domestic activities and lifetime residential history should be collected in future studies.

Dioxin release from illegal cable burning (i.e. the process in which copper and lead are
435 recovered by burning the insulating material from electricity and electronics) may be a critical source of PCDD and PCDF emissions and potentially relevant exposure sources [61]. However, because of the lack of data over the French territory, these emissions could not be considered in the exposure assessment. This might have led to an underestimation of the dioxin exposure and misclassification bias, although non differential, so the risk estimates
440 would be drawn toward unity.

Author's draft

Strengths of our study included its prospective design and the quality of the information collected and regularly updated. The statistical models were adjusted for a large number of individual breast cancer risk factors; therefore, substantial residual confounding is unlikely. Moreover, factors known to influence individual dioxin exposure were considered in the risk estimate, such as age, body mass index, breastfeeding and pregnancy. Human exposure to dioxins can occur from food consumption, inhalation, and occupation. Dietary dioxin intake in relation to breast cancer has been previously investigated in the entire E3N cohort [24]. We also adjusted our models for the dioxin intake and this did not change our results. Moreover, most of the E3N women are teachers and from affiliated occupations, thus their occupational dioxin exposure was assumed to be negligible and homogenous among the study participants.

Focusing our analysis on the E3N women that had permanently lived in the Rhône-Alpes region over the study period restricted the number of participants and thus the power of our study. Confirmation of our findings is required in further studies and larger populations. In particular, further enlargement to breast cancer cases and their matched controls of the entire E3N cohort and identification of dioxin emitting sources at the national level would be relevant.

Conclusion

460

In conclusion, no increase in breast cancer risk was associated with airborne dioxin exposure among women from the E3N cohort. A decrease in breast cancer risk was suggested among subgroups of women, but these results need confirmation in larger populations. The GIS-based metric developed to assess long-term and low-dose airborne dioxin exposure at the individual address level may provide an alternative in absence of measurement of ambient dioxin concentrations. Finally, our study may provide new ways of considering temporal dimensions for environmental exposures in disease risk assessment.

465

List of abbreviations

- 470 AIC, Akaike information criteria
AhR, aryl hydrocarbon receptor
BC, breast cancer
BMI, body mass index
475 CADD, contributive area for dioxin drift
CI, confidence interval
E3N, Étude Épidémiologique auprès de femmes de l'Éducation Nationale
EPIC, European Prospective Investigation into Cancer and Nutrition
ER, estrogen receptor
480 GIS, Geographical Information System
IARC, International Agency for Research Agency
IGN, National Geographic Institute
MET, metabolic equivalent task
MHT, menopausal hormone therapy
485 MSWI, municipal solid waste incinerator
OR, odds ratio
PAH, polycyclic aromatic hydrocarbon
PCB, polychlorinated biphenyls
PCDD, polychlorinated dibenzo-*para*-dioxin
490 PCDF, polychlorinated dibenzofuran
PR, progesterone receptor
SD, standard deviation
TCDD, 2,3,7,8-tetrachlorodibenzo-*para*-dioxin
TEQ, toxic equivalent

495

Declarations

Competing interests

500 The authors declare that they have no competing interests.

Funding

The E3N cohort is financially supported by the French League Against Cancer, the *Mutuelle Générale de l'Education Nationale*, the *Institut Gustave Roussy*, the *Institut National de la Santé et de la Recherche Médicale*. The present project was supported by funding from the
505 Oncostarter Program of the *Cancéropôle Lyon Auvergne Rhône-Alpes*, the French Environment and Energy Management Agency (Grant No 1306C0031), the regional French Cancer League (71) and the Rhône-Alpes Region. Aurélie Danjou is supported by an academic doctoral fellowship from the *Ecole Doctorale Interdisciplinaire Sciences-Santé*,
510 *Université de Lyon*. Aurélie Danjou also received a mobility grant from the *Cancéropôle Lyon Auvergne Rhône-Alpes*.

Authors' contributions

FCC and MCBR have contributed to the conception and design of the E3N study.
515 BF and LD conceived the study on dioxin exposure and breast cancer risk and participated in its design. FCC, MCBR and LD have made contribution to the data acquisition. AMND, EF, LD and TC contributed to the exposure assessment definition and application. EL and KL developed the modeling statistical method. AMND performed the statistical analyses. AMND, BF, EL, KL and LD have interpreted the results. AMND, BF, FCC, LD and MCBR have
520 made contribution to the revision of the manuscript. All authors have read and approved the submitted manuscript.

Acknowledgments

We gratefully acknowledge the participants for providing data and the practitioners for
525 providing pathology reports. We thank Guillaume Harel for his work on the inventory and characterization of the industrial sources located in the Rhône-Alpes region. We thank the scientific committee of the project for its advice on the exposure assessment.

530 **References**

1. Travis CC, Hattemer-Frey HA. A perspective on dioxin emissions from municipal solid waste incinerators. *Risk Anal. Off. Publ. Soc. Risk Anal.* 1989;9:91–7.
2. Zook D, Rappe C. Environmental sources, distribution, and fate of polychlorinated dibenzodioxins, dibenzofurans, and related organochlorines. Schechter A. *Dioxins and health.* New York: Plenum Press; 1994. p. 80–113.
3. ATSDR. Toxicological profile for Chlorinated Dibenzo-p-Dioxins. US Department of Health and Human Services, Agency for Toxic Substances and Disease Registry, Atlanta, GA. [Internet]. 1998 [cited 2016 Sep 24]. Available from: <http://www.atsdr.cdc.gov/toxprofiles/tp104.pdf>
- 540 4. Anzivino-Viricel L, Falette N, Carretier J, Montestrucq L, Guye O, Philip T, et al. Gestion des déchets ménagers et assimilés: bilan des connaissances et évaluation des effets sanitaires en population générale et au travail. *Environ. Risques Santé.* 2012;11:360–377.
5. Nzihou A, Themelis NJ, Kemiha M, Benhamou Y. Dioxin emissions from municipal solid waste incinerators (MSWIs) in France. *Waste Manag.* 2012;32:2273–7.
- 545 6. CITEPA. Dioxines et Furanes - PCDD-F [Internet]. 2015 [cited 2016 Sep 1]. Available from: <http://www.citepa.org/fr/air-et-climat/polluants/polluant-organiques-persistants/dioxines-et-furanes>
7. International Agency for Research on Cancer, Weltgesundheitsorganisation, editors. IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 F, chemical agents and related occupations: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 20 - 27 October 2009. Lyon: IARC; 2012.
- 550 8. Baan R, Grosse Y, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. A review of human carcinogens--Part F: chemical agents and related occupations. *Lancet Oncol.* 2009;10:1143–4.
- 555 9. Jenkins S, Betancourt AM, Wang J, Lamartiniere CA. Endocrine-active chemicals in mammary cancer causation and prevention. *J. Steroid Biochem. Mol. Biol.* 2012;129:191–200.
10. Birnbaum LS, Fenton SE. Cancer and Developmental Exposure to Endocrine Disruptors. *Environ. Health Perspect.* 2002;111:389–94.
- 560 11. Mandal PK. Dioxin: a review of its environmental effects and its aryl hydrocarbon receptor biology. *J. Comp. Physiol. B.* 2005;175:221–30.
12. Bertazzi A, Pesatori AC, Consonni D, Tironi A, Landi MT, Zocchetti C. Cancer incidence in a population accidentally exposed to 2,3,7,8-tetrachlorodibenzo-para-dioxin. *Epidemiol. Camb. Mass.* 1993;4:398–406.
- 565 13. Pesatori A, Consonni D, Rubagotti M, Grillo P, Bertazzi P. Cancer incidence in the population exposed to dioxin after the “Seveso accident”: twenty years of follow-up. *Environ. Health.* 2009;8:39.

Author's draft

- 570 14. Warner M, Mocarelli P, Samuels S, Needham L, Brambilla P, Eskenazi B. Dioxin exposure and cancer risk in the Seveso Women's Health Study. *Environ. Health Perspect.* 2011;119:1700–5.
15. Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. Environmental pollutants and breast cancer: Epidemiologic studies. *Cancer.* 2007;109:2667–711.
- 575 16. Manuwald U, Garrido MV, Berger J, Manz A, Baur X. Mortality study of chemical workers exposed to dioxins: follow-up 23 years after chemical plant closure. *Occup. Environ. Med.* 2012;69:636–642.
17. Revich B, Aksel E, Ushakova T, Ivanova I, Zhuchenko N, Klyuev N, et al. Dioxin exposure and public health in Chapaevsk, Russia. *Chemosphere.* 2001;43:951–966.
- 580 18. Fabre P, Daniau C, Gorla S, de Crouy-Chanel P, Empereur-Bissonnet P. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. *Synthèse St.-Maurice InVS* [Internet]. 2008 [cited 2015 Aug 21]; Available from: http://www.le1000gru.org/ital/2009/scrfiles/rapport_uiom.pdf
- 585 19. Viel J-F, Clément M-C, Hägi M, Grandjean S, Challier B, Danzon A. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure. *Int. J. Health Geogr.* 2008;7:4.
20. Cordioli M, Ranzi A, De Leo GA, Lauriola P. A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators. *J. Environ. Public Health.* 2013;2013:1–12.
- 590 21. Vacek PM. Assessing the effect of intensity when exposure varies over time. *Stat. Med.* 1997;16:505–13.
22. Lamartiniere CA. Timing of exposure and mammary cancer risk. *J. Mammary Gland Biol. Neoplasia.* 2002;7:67–76.
23. Teitelbaum SL, Belpoggi F, Reinlib L. Advancing research on endocrine disrupting chemicals in breast cancer: Expert panel recommendations. *Reprod. Toxicol.* 2015;54:141–7.
- 595 24. Danjou A, Fervers B, Boutron-Ruault M-C, Philip T, Clavel-Chapelon F, Dossus L. Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort. *Breast Cancer Res.* 2015;17:39.
25. Clavel-Chapelon F, E3N Study Group. Cohort Profile: The French E3N Cohort Study. *Int. J. Epidemiol.* 2015;44:801–9.
- 600 26. Binachon B, Dossus L, Danjou AMN, Clavel-Chapelon F, Fervers B. Life in urban areas and breast cancer risk in the French E3N cohort. *Eur. J. Epidemiol.* 2014;29:743–51.
27. Riboli E. Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann. Oncol.* 1992;3:783–791.
- 605 28. Riboli E, Hunt K, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr.* 2002;5:1113.

Author's draft

29. UNEP (United Nations Environment Programme). Toolkit for Identification and Quantification of Releases of Dioxins, Furans and Other Unintentional POPs [Internet]. 2013 [cited 2016 Aug 22]. Available from: http://toolkit.pops.int/Publish/Main/01_Index.html
- 610 30. Hoek G, Fischer P, Van Den Brandt P, Goldbohm S, Brunekreef B. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J. Expo. Anal. Environ. Epidemiol.* [Internet]. 2001 [cited 2016 Aug 12];11.
31. Yu C-L. Residential Exposure to Petrochemicals and the Risk of Leukemia: Using Geographic Information System Tools to Estimate Individual-Level Residential Exposure. *Am. J. Epidemiol.* 2006;164:200–7.
- 615 32. Vienneau D, de Hoogh K, Briggs D. A GIS-based method for modelling air pollution exposures across Europe. *Sci. Total Environ.* 2009;408:255–66.
33. White N, teWaterNaude J, van der Walt A, Ravenscroft G, Roberts W, Ehrlich R. Meteorologically estimated exposure but not distance predicts asthma symptoms in schoolchildren in the environs of a petrochemical refinery: a cross-sectional study. *Environ. Health* [Internet]. 2009 [cited 2016 Aug 12];8.
- 620 34. Gulliver J, Briggs D. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. *Sci. Total Environ.* 2011;409:2419–29.
35. Pronk A, Nuckols JR, De Roos AJ, Airola M, Colt JS, Cerhan JR, et al. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case-control study. *Environ. Health.* 2013;12:1.
- 625 36. Soulhac L, Salizzoni P, Cierco F-X, Perkins R. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmos. Environ.* 2011;45:7379–95.
37. Soulhac L, Salizzoni P, Mejean P, Didier D, Rios I. The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmos. Environ.* 2012;49:320–37.
- 630 38. Viera AJ, Garrett JM, others. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 2005;37:360–363.
39. Tehard B, Lahmann PH, Riboli E, Clavel-Chapelon F, the E3N Group. Anthropometry, breast cancer and menopausal status: Use of repeated measurements over 10 years of follow-up? results of the French E3N women's cohort study. *Int. J. Cancer.* 2004;111:264–9.
- 635 40. Uemura H, Arisawa K, Hiyoshi M, Satoh H, Sumiyoshi Y, Morinaga K, et al. PCDDs/PCDFs and dioxin-like PCBs: Recent body burden levels and their determinants among general inhabitants in Japan. *Chemosphere.* 2008;73:30–7.
41. Humblet O, Williams PL, Korrick SA, Sergeev O, Emond C, Birnbaum LS, et al. Predictors of Serum Dioxin, Furan and PCB Concentrations among Women from Chapaevsk, Russia. *Environ. Sci. Technol.* 2010;44:5633–40.
- 640 42. Caspersen IH, Knutsen HK, Brantsæter AL, Haugen M, Alexander J, Meltzer HM, et al. Dietary exposure to dioxins and PCBs in a large cohort of pregnant women: Results from the Norwegian Mother and Child Cohort Study (MoBa). *Environ. Int.* 2013;59:398–407.
- 645

Author's draft

43. Wang RY, Jain RB, Wolkin AF, Rubin CH, Needham LL. Serum Concentrations of Selected Persistent Organic Pollutants in a Sample of Pregnant Females and Changes in Their Concentrations during Gestation. *Environ. Health Perspect.* 2009;117:1244–9.
- 650 44. Hauptmann M, Wellmann J, Lubin JH, Rosenberg PS, Kreienbrock L. Analysis of Exposure-Time-Response Relationships Using a Spline Weight Function. *Biometrics.* 2000;56:1105–1108.
45. Warner M, Eskenazi B, Mocarelli P, Gerthoux PM, Samuels S, Needham L, et al. Serum dioxin concentrations and breast cancer risk in the Seveso Women's Health Study. *Environ. Health Perspect.* 2002;110:625.
- 655 46. Raaschou-Nielsen O, Pavuk M, LeBlanc A, Dumas P, Weber JP, Olsen A, et al. Adipose organochlorine concentrations and risk of breast cancer among postmenopausal Danish women. *Cancer Epidemiol. Biomarkers Prev.* 2005;14:67–74.
47. Gammon MD, Wolff MS, Neugut AI, Eng SM, Teitelbaum SL, Britton JA, et al. Environmental toxins and breast cancer on Long Island. II. Organochlorine compound levels in blood. *Cancer Epidemiol. Biomarkers Prev.* 2002;11:686–697.
- 660 48. Rusiecki JA, Holford TR, Zahm SH, Zheng T. Polychlorinated biphenyls and breast cancer risk by combined estrogen and progesterone receptor status. *Eur. J. Epidemiol.* 2004;19:793–801.
49. Iwasaki M, Inoue M, Sasazuki S, Kurahashi N, Itoh H, Usuda M, et al. Plasma organochlorine levels and subsequent risk of breast cancer among Japanese women: A nested case-control study. *Sci. Total Environ.* 2008;402:176–83.
- 665 50. Itoh H, Iwasaki M, Hanaoka T, Kasuga Y, Yokoyama S, Onuma H, et al. Serum organochlorines and breast cancer risk in Japanese women: a case-control study. *Cancer Causes Control CCC.* 2009;20:567–80.
- 670 51. Wang WL, Porter W, Burghardt R, Safe SH. Mechanism of inhibition of MDA-MB-468 breast cancer cell growth by 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin. *Carcinogenesis.* 1997;18:925–933.
52. Zhang S, Lei P, Liu X, Li X, Walker K, Kotha L, et al. The aryl hydrocarbon receptor as a target for estrogen receptor-negative breast cancer chemotherapy. *Endocr. Relat. Cancer.* 2009;16:835–44.
- 675 53. Yoshioka H, Hiromori Y, Aoki A, Kimura T, Fujii-Kuriyama Y, Nagase H, et al. Possible aryl hydrocarbon receptor-independent pathway of 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced antiproliferative response in human breast cancer cells. *Toxicol. Lett.* 2012;211:257–65.
- 680 54. Kociba RJ, Keyes DG, Beyer JE, Carreon RM, Wade CE, Dittenber DA, et al. Results of a two-year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in rats. *Toxicol. Appl. Pharmacol.* 1978;46:279–303.
- 685 55. Collins JJ, Bodner K, Burns CJ, Budinsky RA, Lamparski LL, Wilken M, et al. Body mass index and serum chlorinated dibenzo-p-dioxin and dibenzofuran levels. *Chemosphere.* 2007;66:1079–85.

Author's draft

56. Emond C, DeVito M, Warner M, Eskenazi B, Mocarelli P, Birnbaum LS. An assessment of dioxin exposure across gestation and lactation using a PBPK model and new data from Seveso. *Environ. Int.* 2016;92–93:23–32.

690 57. Moon H-B, Lee D-H, Lee YS, Kannan K. Concentrations and accumulation profiles of PCDDs, PCDFs and dioxin-like PCBs in adipose fat tissues of Korean women. *J. Environ. Monit.* 2011;13:1096.

58. Dhooze W, Den Hond E, Koppen G, Bruckers L, Nelen V, Van De Mierop E, et al. Internal exposure to pollutants and body size in Flemish adolescents and adults: Associations and dose–response relationships☆. *Environ. Int.* 2010;36:330–7.

695 59. White AJ, Teitelbaum SL, Stellman SD, Beyea J, Steck SE, Mordukhovich I, et al. Indoor air pollution exposure from use of indoor stoves and fireplaces in association with breast cancer: a case-control study. *Environ. Health.* 2014;13:1.

700 60. Basagaña X, Aguilera I, Rivera M, Agis D, Foraster M, Marrugat J, et al. Measurement Error in Epidemiologic Studies of Air Pollution Based on Land-Use Regression Models. *Am. J. Epidemiol.* 2013;178:1342–6.

61. Stockholm Convention Clearing House. Section VI. Guidance/guidelines by source category: Part III of Annex C. VI.L. Smouldering of copper cables [Internet]. [cited 2016 Oct 7]. Available from:
705 <http://chm.pops.int/Implementation/BATandBEP/BATBEPGuidelinesArticle5/tabid/187/Default.aspx>

Tables

710 Table 1. Baseline characteristics of invasive breast cancer cases and matched controls, N=1215.

	Cases N=429	Controls N=786	P ^a
Age (years), mean (SD)	51.7 (7.1)	51.2 (6.9)	0.18
Age at breast cancer diagnosis (years), mean (SD)	58.0 (7.8)	-	-
Menopausal status, n (%)			0.30
Premenopausal	212 (49.4)	413 (52.5)	
Postmenopausal	217 (50.6)	373 (47.5)	
Body Mass Index (kg/m ²), n (%)			0.78
< 25	366 (85.3)	665 (84.6)	
[25-30[53 (12.4)	97 (12.3)	
≥30	10 (2.3)	24 (3.1)	
Alcohol intake, n (%)			0.01
0 g/day	37 (8.6)	97 (12.3)	
< 6.1 g/day	146 (34.0)	256 (32.6)	
≥ 6.1 g/day	215 (50.1)	340 (43.3)	
missing data	31 (7.2)	93 (11.8)	
Smoking status, n (%)			0.30
Never smoker	235 (54.8)	446 (56.7)	
Ex-smoker	55 (12.8)	114 (14.5)	
Current smoker	139 (32.4)	226 (28.8)	
Age at menarche (years), n (%)			0.30
< 12	89 (20.8)	146 (15.6)	
[12-14[231 (53.9)	396 (50.4)	
≥ 14	109 (25.4)	244 (31.0)	
Previous use of oral contraceptives, n (%)	264 (61.5)	469 (59.7)	0.60
Previous use of progestin before menopause, n (%)	200 (46.6)	339 (43.1)	0.21
Parity and age at first full-term pregnancy, n (%)			0.02
≤ 2 children and age at first birth < 30	197 (45.9)	387 (49.2)	
≤ 2 children and age at first birth ≥ 30	63 (14.7)	67 (8.5)	
≥ 3 children	109 (25.4)	235 (29.9)	
missing data	60 (14.0)	97 (12.3)	
Breastfeeding, n (%)	243 (55.6)	479 (60.9)	0.20
Previous family history of breast cancer, n (%)	76 (17.7)	78 (9.9)	<0.0001
Previous history of personal benign breast disease, n (%)	166 (38.7)	228 (29.0)	0.0004
Estimated dietary dioxin intake (pg-TEQ/kg body weight/day), mean (SD)	1.3 (0.4)	1.2 (0.4)	0.25
missing data	31 (7.2)	93 (11.8)	

^a conditional logistic regression, except for age and menopausal status, which were matched factors (unconditional logistic regression)

Table 2. Odds ratios (OR) for the association between invasive breast cancer and cumulative airborne dioxin exposure, and according to hormone receptor status and menopausal status (N=1215), 1990-2008.

Populations	Cumulative airborne dioxin exposure ($\mu\text{g-TEQ}/\text{m}^2$)							
	/ 1 SD ^a	[0.0, 2.4E-5]	[2.4E-5, 2.6E-3]	[2.6E-3, 2.1E-2]	[2.1E-2, 1.0E-1]	[1.0E-1, 24.4]	P-trend	
Overall								
N cases	429	71	97	91	96	74		
OR (95% CI) ^b	0.99 (0.87-1.13)	1.00	1.23 (0.81-1.87)	1.17 (0.75-1.81)	1.05 (0.68-1.65)	0.76 (0.47-1.23)		0.18
Premenopause								
N cases	212	33	52	39	49	39		
OR (95% CI) ^b	1.04 (0.89-1.22)	1.00	1.44 (0.76-2.71)	1.47 (0.72-3.02)	1.45 (0.69-3.04)	1.05 (0.48-2.29)		0.94
Postmenopause								
N cases	217	38	45	52	47	35		
OR (95% CI) ^b	0.55 (0.26-1.14)	1.00	1.46 (0.63-3.39)	1.44 (0.66-3.16)	1.07 (0.47-2.48)	0.94 (0.41-2.19)		0.68
ER positive^c								
N cases	244	38	57	51	58	40		
OR (95% CI) ^b	1.01 (0.88-1.15)	1.00	1.56 (0.93-2.63)	1.39 (0.80-2.43)	1.31 (0.76-2.28)	0.80 (0.44-1.47)		0.34
ER negative^c								
N cases	80	17	11	17	19	16		
OR (95% CI) ^b	0.85 (0.54-1.32)	1.00	0.56 (0.24-1.31)	0.83 (0.36-1.90)	0.87 (0.38-1.98)	0.70 (0.29-1.69)		0.73
PR positive^d								
N cases	215	38	44	44	53	36		
OR (95% CI) ^b	1.02 (0.90-1.17)	1.00	1.08 (0.63-1.86)	0.98 (0.56-1.73)	1.03 (0.59-1.80)	0.61 (0.33-1.12)		0.14
PR negative^d								
N cases	108	17	24	22	24	21		
OR (95% CI) ^b	0.80 (0.50-1.26)	1.00	1.48 (0.71-3.08)	1.53 (0.69-3.362)	1.39 (0.60-3.04)	1.28 (0.55-2.98)		0.68

^a 1 SD=0.88 $\mu\text{g-TEQ}/\text{m}^2$

Author's draft

720 ^b adjusted for height, body mass index, family history of breast cancer, personal history of benign breast disease, mammography before baseline, use of oral contraceptives, use of oral progestatives, smoking status, breastfeeding, education, physical activity, alcohol consumption, age at menarche, parity and age at first full term pregnancy, use of menopausal hormone therapy and age at menopause among postmenopausal women, and status of birthplace

^c n=105 invasive breast cancer cases with missing estrogen receptor status

^d n=106 invasive breast cancer cases with missing progesterone receptor status

725

Table 3. Odds ratios (OR) for the association between invasive breast cancer and cumulative airborne dioxin exposure, stratified by BMI, age at first full-term pregnancy and breastfeeding (N=1215), 1990-2008.

Populations	Cumulative airborne dioxin exposure ($\mu\text{g-TEQ}/\text{m}^2$)								P-trend
	/ 1 SD ^a	[0.0, 2.4E-5]	[2.4E-5, 2.6E-3]	[2.6E-3, 2.1E-2]	[2.1E-2, 1.0E-1]	[1.0E-1, 24.4]			
BMI \leq 21.8 kg/m² ^c		38	52	48	50	33			
N cases	221	1.00	1.23 (0.68-2.25)	1.12 (0.58-2.16)	0.78 (0.41-1.48)	0.49 (0.24-0.98)			0.01
OR (95% CI) ^b	1.07 (0.91-1.25)								
BMI > 21.8 kg/m² ^c		33	45	43	46	41			
N cases	208	1.00	1.25 (0.68-2.30)	1.26 (0.66-2.41)	1.46 (0.75-2.86)	1.10 (0.54-2.25)			0.70
OR (95% CI) ^b	0.75 (0.50-1.11)								
Age at FFTP \leq 24 years^d		33	38	41	36	13			
N cases	161	1.00	0.70 (0.35-1.36)	0.99 (0.49-2.01)	0.52 (0.25-1.08)	0.15 (0.06-0.37)			0.0003
OR (95% CI) ^b	0.24 (0.05-1.12)								
Age at FFTP > 24 years^d		32	52	39	48	37			
N cases	208	1.00	1.74 (0.94-3.22)	1.28 (0.66-2.47)	1.64 (0.83-3.26)	1.13 (0.56-2.29)			0.92
OR (95% CI) ^b	1.10 (0.92-1.33)								
Ever breastfed		41	59	51	56	36			
N cases	243	1.00	0.98 (0.56-1.70)	1.05 (0.59-1.89)	0.96 (0.53-1.73)	0.41 (0.22-0.79)			0.01
OR (95% CI) ^b	1.01 (0.89-1.15)								
Never breastfed		30	38	40	40	38			
N cases	186	1.00	1.83 (0.92-3.62)	1.27 (0.61-2.65)	1.20 (0.57-2.55)	1.81 (0.81-4.02)			0.44
OR (95% CI) ^b	0.81 (0.51-1.26)								

^a 1 SD=0.88 $\mu\text{g-TEQ}/\text{m}^2$

^b adjusted for height, body mass index, family history of breast cancer, personal history of benign breast disease, mammography before baseline, use of oral contraceptives, use of oral progestatives, smoking status, breastfeeding, education, physical activity, alcohol consumption, age at

Author's draft

menarche, parity and age at first full term pregnancy, use of menopausal hormone therapy and age at menopause among postmenopausal women, and status of birthplace

^c median BMI based on the distribution among controls

^d median age at first full-term pregnancy based on the distribution among controls

Author's draft

Table 4. Odds ratios (OR) for the association between invasive breast cancer and cumulative airborne dioxin exposure according to time before diagnosis, 1990-2008.

Time before diagnosis (years)	Cumulative airborne dioxin exposure (/0.1 µg-TEQ/m ²)			
	0	5	10	15
N cases	211	279	166	55
OR (95%CI) ^a	1.00 (0.88-1.02)	1.00 (0.92-1.02)	1.00 (0.89-1.01)	1.00 (0.84-1.01)

740 ^a adjusted for matching factors and status of birthplace

3. Additional analysis

Analyses regarding the association between airborne dioxin exposure and breast cancer risk among E3N women of the Rhône-Alpes region are still ongoing. Along with the elaboration of the methodology, the team worked on a new questionnaire on lifetime residential history, to be administered to the E3N women. These aspects are developed below.

3.1. Time-dependent risk estimation

Results presented in Article #3 regarding the estimation of time-dependent odds ratios are preliminary. First models included matching factors and birthplace status, and B-splines were estimated with one knot. In future analyses, models will be specified with varying number of knots (one, two or three) and will include other breast cancer risk factors. The final model will be chosen in order to minimize the Akaike Information Criterion.

3.2. Questionnaire on Lifetime Residential History

In our study, the residential addresses of the participants are lacking before entering the cohort in 1990 and information on past dioxin emissions from industries prior to 1990 are extremely sparse. Residence in urban places has been suggested to be a surrogate for air pollution exposure released from road traffic, industrial facilities and waste incineration (Lambrechtsen et al. 2012). We previously investigated the association between birth or residence in an urban area and breast cancer risk among 75,889 participants of the E3N cohort. Being born in an urban area was associated with increased breast cancer risk (HR=1.07, 95%CI=1.01-1.14) (Binachon et al. 2014) (Appendice 6). In order to allow adjustment for earlier exposures, based on the rural/urban status of the place of residence as a surrogate for exposures before 1990, a supplementary self-administered questionnaire, named Questionnaire on Residential History (QRH), was designed by the team project and sent to the women of the GEO3N study population living in the Rhône-Alpes region; the QRH is presented in Appendice 4. Quality and quantity of data collected were analyzed prior to applying the QRH to obtain the lifetime residential history from the participants (not deceased) of the nested case-control study at the national level.

3.2.1. Description of the QRH

The QRH first asked women to complete their lifetime residential history, from one year before birth to 2014. Street address (number, name and additional information such as the floor), zip code, town or city, country and residency periods (starting and ending years) were recorded to be geolocalized in the GIS. A remarkable point (*e.g.* the church, town hall, river etc.) could be indicated when the address had been forgotten. Women could specify up to 20 addresses as well as the address occupied during their mother's pregnancy. The questionnaire also aimed at collecting characteristics of the successive homes, such as domestic heating (collective central heating; heating with wood or coal; individual heating with fuel, gas and electricity), combustion activities (use of a stove or a wood stove; use of an open fireplace; use of a closed fireplace; barbecue cooking; burning of green waste) and eating habits (consumption frequency of fruits and vegetables, eggs, dairy products, meat and seafood, produced locally, *i.e.* within 2 km from their residence). A second part of the QRH aimed at collecting lifetime addresses of their places of education and work (name of the school or the workplace; street address; starting and ending years). For each of these places, women could indicate the time spent by day on this place, the journey time from home to work and the modes of transport used (on foot; bike; motorized wheels; diesel vehicle; gasoline vehicle; public transportation).

The QRH included many questions, repeated for each address of residence and workplace, and relied on the memory of women born between 1925 and 1950 (*i.e.* aged 65 to 90 years at the time of sending). Filling errors, omissions and incomplete answers or questionnaires were therefore expected from the project team. Being aware of the potential difficulty for the E3N women to complete the QRH, it was first sent to the women living in the Rhône-Alpes region.

The elaboration and sending of the questionnaire were carried out by the E3N team in Villejuif. The QRH was first sent on February 2014 to 1,642 women of the GEO3N study population of the Rhône-Alpes region that were not deceased. One reminder was made for the non-respondents, in June 2014, by mail. The data cleaning of the QRH began in May 2015, after 1,027 questionnaires were returned. Data were imported into a database from the optical reading of the questionnaires. All of the addresses recorded were entered manually by a partner company (Vivetic). Data were cleaned and analysed to determine quality and quantity of data that could be used to clarify the estimated exposure to dioxins in the Rhône-Alpes region, and consider whether or not its distribution to the entire study population of GEO3N (excluding the ones from the Rhône-Alpes region).

3.2.2. Analyses of the QRH

Results regarding quality and content of the filling of the QRH are presented in Appendice 5. Briefly, the QRH had a 68% return rate (1,027 returns/1,642 mailings). The response rate was 57% after one reminder with 973 and 977 questionnaires on residential and occupational histories that could be analyzed. Errors in the filling of the QRH referred to non-chronological residency periods and global non-understanding for some questions. Missing data were frequent for earlier periods of life and for the consumption frequency of food produced locally. The QRH will be sent to all participants of the GEO3N, after simplification of some of the questions.

3.2.3. Geocoding of the lifetime residential and occupational addresses

Lifetime residential and occupational addresses were geocoded and their accuracy was evaluated: 78.7% of addresses were geocoded at the address or at the street segment (Appendice 5). A “rural” or “urban” status was assigned to each address according to the census defined by The National Institute of Statistics and Economic Studies (INSEE). An urban commune was defined as a commune with more than 2,000 inhabitants and with buildings separated by less than 200 m; otherwise the commune was defined as rural (Insee 2016). This status will be used to compute the number of years of living in an urban area from birth to inclusion in the E3N cohort that will be considered as a surrogate for airborne dioxin exposure to account for past exposures in further statistical analysis.

4. Conclusion

In the investigation on the association between cumulative airborne dioxin exposure and breast cancer risk, we did not observe any increase in breast cancer risk overall, nor among pre- and postmenopausal women and for hormone-receptor defined breast tumors. Further adjustment on estimated dietary dioxin exposure did not change our results (Article #1). Inverse associations between cumulative airborne dioxin and breast cancer risk were found among women with low BMI, women young at maternal age and women with a history of breastfeeding. However, due to the small number of cases in subgroup analysis and multiple comparisons, our results need to be confirmed in a larger population.

The nested case-control study design along with an incidence density sampling allowed reducing the size of the study population while remaining in a prospective context. Moreover, breast cancer cases and controls were matched on age and date of blood donation or inclusion in the E3N cohort so

that the case-control pair shared the same window of exposure.

Although the association between airborne dioxin exposure and breast cancer risk was investigated within a case-control study, we could consider the timing of exposure in statistical analysis, by assigning weights to past exposures. First results did not show any impact of cumulative airborne dioxin exposure according to time before diagnosis on the risk of breast cancer. However, these were preliminary results and the estimation of time-dependent odds ratios for breast cancer needs further analysis.

The inventory of dioxin emitting sources was non-exhaustive and focused on major industrial sources, not considering domestic activities and traffic. Although the latter poorly contributed to airborne dioxin exposure, this might have led to an underestimation of exposure. Moreover, we could not consider dioxin emissions from hospital waste incinerators at the time of analysis; however, they will be accounted for in further analysis.

A major limitation included the lack of knowledge about the participants' dioxin exposure before entering the cohort in 1990; therefore, our estimation of their airborne dioxin exposure might be underestimated, as dioxins are known to accumulate and have a long half-life in the human body. However, we designed a new questionnaire retracing the lifetime residential and occupational histories of the participants living in the Rhône-Alpes region, and additional information was collected for each address regarding personal heating habits (chimney fires, stove, etc.), personal burning of green waste and eating of locally produced food. The information recorded will allow to assess domestic dioxin exposure and to compute a proxy for dioxin exposure from birth to 1990 using the "rural" or "urban" status for each participant address. However, the information in the Rhône-Alpes region was available only for a small number of subjects and its analysis might lack power. The questionnaire will be modified and sent to all women from the GEO3N study population.

This work was restricted to women that had lived permanently in the Rhône-Alpes region in order to elaborate the methodology for the assessment of airborne dioxin exposure and the estimation of breast cancer risk. The study is now extended to the national French territory, among nearly 5,500 breast cancer cases and 5,500 matched controls.

CHAPTER V. DISCUSSION

1. Discussion

My doctoral research aimed at investigating the association between dioxin exposure and breast cancer risk, providing new knowledge about historical long-term and multi-source dioxin exposure in the general population and filling the methodological gaps of the published literature regarding environmental exposure assessment and the timing of exposure in the risk estimation. The first objective was to examine dietary dioxin exposure in relation to breast cancer risk in the E3N cohort (Danjou et al. 2015). The second objective was to develop a GIS-based metric for the assessment of airborne dioxin exposure (Faure et al., submitted). The third objective was to examine airborne dioxin exposure in relation to breast cancer risk in a case-control study nested within the E3N cohort (Danjou, in preparation).

We were the first, to our knowledge, to consider multiple exposure routes and to investigate dietary dioxin exposure and breast cancer risk. Furthermore, in comparison to previous studies, we were the first to assess multiple sources contributing to airborne dioxin exposure. To overcome methodological limitations of published studies, our research has focused on the development of a GIS-based metric to assess airborne dioxin exposure.

Overall, no association with dioxin exposure was observed among women from the E3N cohort, except for a decrease in hormone-independent breast cancer risk. The observed decreased risk was statistically significant among postmenopausal women with the highest dioxin intake (Danjou et al. 2015). Also, although they were not statistically significant, our results suggest that low-dose cumulative airborne dioxin exposure may be associated with decreased ER-negative breast cancer risk. The observed inverse association is supported by *in vitro* evidence and animal studies demonstrating the anti-proliferative action of dioxins (Kociba et al. 1978; Wang et al. 1997; Yoshioka et al. 2012; Zhang et al. 2009). Moreover, these findings are consistent with results from epidemiological studies on organochlorine pesticides and PCBs (Gammon et al. 2002b; Raaschou-Nielsen et al. 2005). Finally, a decreased breast cancer risk was observed among women over 60 years of age and most exposed to emissions from a French MSWI; however, breast cancer risk was not estimated according to hormone-receptor status (ER, PR) (Viel et al. 2008).

Furthermore, we observed a decrease in breast cancer risk associated with cumulative airborne dioxin exposure among subgroups defined by low BMI, younger maternal age and a history of breastfeeding. However, mechanisms underlying these effects remain not completely understood, and results might have been due to chance. Regarding dietary dioxin exposure, no statistically significant association with breast cancer was observed in subgroups defined by BMI and breastfeeding. Subgroup analyses according to specific windows of breast vulnerability were not performed and the association between dietary dioxin exposure and breast cancer risk should be further analyzed according to age at menarche and age at first full-term pregnancy.

While adjustment for dietary dioxin exposure in the GIS-based assessment did not have an impact on risk estimates, it would be pertinent to further assess dioxin exposure by considering the risk associated with both dietary and airborne exposures in an integrated model to investigate the joint effect of these exposure routes (Petit et al. 2012).

Two concepts have arisen regarding exposure to endocrine disrupting chemicals, including dioxins. First, effects of EDCs might occur even at low-doses, *i.e.* in the range of human exposures to environmentally relevant doses or below doses traditionally investigated in toxicological studies. Although mean serum TCDD levels have decreased this past forty years, animal and experimental studies suggest that even low levels of dioxins could have adverse effects on health outcomes (Vandenberg et al. 2012). Second, the relationship between EDC exposure and health outcomes might not be linear but non-monotonic. A dose-response curve is non-monotonic when the slope of the curve changes sign within the range of exposures examined (Vandenberg et al. 2012). In our study

population, we observed low-dose effects of dioxins on the risk of breast cancer, for both dietary and airborne exposures, and with non-monotonic dose-response curves. However, they were not statistically significant. One reason to explain the non-statistically significant association between airborne dioxin exposure and breast cancer risk is the lack power in the regional study population. A second reason is the non-investigation of non-monotonic relationship between both dietary and airborne exposures and breast cancer risk. Although, the hypothesis of a low-dose non-monotonic dose-response of EDCs on health outcomes in humans has been questioned (Kamrin 2007; Rhomberg and Goodman 2012), new methods have been developed to quantitatively assess non-monotonic dose-response relationships for the risk assessment (Lagarde et al. 2015), and future research should investigate potential non-monotonic curves in the risk assessment.

Strengths of our study include the E3N prospective cohort, which has prospectively collected quality information on health status, lifestyle and reproductive factors to support the study of many public health issues, including breast cancer. E3N involves a large number of participants, with a long-term follow-up (since 1990) and regular updates of their personal and health information that allowed the prospective identification of breast cancer cases, the collection of food consumption data and the reconstruction of the residential history over the study period. We were able to assess both dietary and airborne dioxin exposures within the E3N cohort, which are the two main routes for dioxin exposure. We developed a GIS-based metric including parameters that were identified through a literature review and integrated in the GIS. Several combinations of these parameters were tested by comparison with dispersion modeling and the final metric included parameter combinations that showed the best agreement between GIS and dispersion modeling methods. The definition of the GIS-based metric was confirmed by a validation in another area and at different periods. Specific statistical models were applied in order to estimate breast cancer risk according to time before diagnosis.

Although our research project improved the methodology regarding the study of dioxin exposure and breast cancer risk, some aspects could not be considered: due to limitations and difficulties in the exposure assessment; measurement of dioxin concentrations in blood or adipose tissue; exposures occurring early in life; lifetime exposure; and combined exposure.

We faced several difficulties when assessing both dietary and airborne dioxin exposures. The estimated dietary dioxin exposure did not consider the origin of the food consumed, resulting in an underestimation of the dioxin intake of women living in the vicinity of industrial sources who could be exposed to food locally produced and more contaminated than average. The assessment of airborne dioxin exposure comprised several steps, some of which had limitations. Although the inventory and characterization of the dioxin emitting sources were necessary in order to identify the different industrial sectors and assigned them dioxin emission intensity, the data collection was time consuming and requests made directly to facilities were unsuccessful for most. Moreover, dioxin exposure from domestic activities, traffic and from illegal activities were not considered in the assessment and could have led to an underestimation of dioxin exposure among participants exposed to these extra-sources. The participants' residential history was extracted from past E3N questionnaires and was not initially recorded to be geocoded; this could have led to a potential misclassification bias due to poor positional accuracy. However, we examined the positional accuracy of the automatic geocoding method: most of addresses were geocoded to the exact address or the street segment, and the median positional error compared to a manual geocoding was small regarding the size of the impact zone of dioxin emissions. The impact of the positional error on exposure assessment and risk estimation should be further investigated through sensitivity analyses. Our GIS-based metric was elaborated from a literature review that allowed the identification of parameters known to influence airborne dioxin exposure (Gulliver and Briggs 2011; Hoek et al. 2001; Pronk et al. 2013; Vienneau et al. 2009; White et al. 2009; Yu 2006). Although we improved simple GIS-proximity models by adding parameters related to meteorology, topography and characteristics of industrial sources, other parameters, such as

atmospheric turbulence, ambient temperature and traffic-related factors, and the use of other exposure assessment methods, such as air dispersion modeling and land use regression models, might have further improved exposure assessment and produced more accurate exposure estimates (Zou et al. 2009). However, we compared our GIS-based estimation of airborne dioxin exposure with dioxin concentrations from a dispersion modeling, applied in two restricted areas, and agreement between these two methods ranged from “substantial” to “almost perfect”. Moreover, because dioxin emission monitoring data were sparse in the French territory and decreased over the study period, we could not use dispersion modeling or LUR to assess airborne dioxin exposure in the E3N cohort. These limitations on exposure assessment might have led to non-differential misclassification bias and risk estimates drawn toward unity.

In the study of the association between dietary dioxin exposure and breast cancer risk, the 15-year follow-up and the size of the population allowed a statistical power enough to show low risk estimates. We cannot exclude a lack of statistical power regarding analyses performed among women from the nested case-control study restricted to the Rhône-Alpes region, and interpretation of results were limited due to the sample size. However, the study in this restricted area aimed at developing and implementing the methodology regarding exposure assessment and risk estimation. The study of the association between airborne dioxin exposure and breast cancer risk among the entire GEO3N study population (5,500 breast cancer cases and 5,500 matched controls) will be well powered to estimate the association between airborne dioxin exposure and breast cancer risk.

In our research project, dioxin exposure was not assessed with direct measures obtained from personal monitoring or biological measurements. Although these measurement methods can be considered as the best measures for assessing the effect of a substance, they do not provide a “gold standard”, because regarding dioxins, they may not reflect the cumulative and historical exposure – except for measurement in adipose tissue, and are conditioned by metabolism, age, adiposity, breastfeeding and dietary habits. Moreover, measurement of dioxins in human biological tissues, including blood (plasma), adipose tissue, hair and saliva, faces several practical difficulties. Measurements are expensive and biological samples must be available in sufficient quantity. The sampling of subcutaneous or abdominal adipose tissue is an invasive procedure, difficult to implement in epidemiological studies or possibly resulting in small sample sizes with poor statistical power (Brody and Rudel 2003). No conclusive results were obtained from dioxin measurements in hair and saliva (Nakao et al. 2005; Ogawa et al. 2003). Estimations of both dietary and airborne dioxin exposures were not compared to direct measures. In the E3N cohort, participants provided blood samples in 1995-1997; however, quantity available was not enough for dioxins to be measured. With the aim of correlating the estimated dietary and airborne dioxin exposures with measured dioxin exposure, we considered conducting a new biological sampling (50 to 100 mL of blood) among controls residing in the Rhône-Alpes region, and for whom both dietary and airborne dioxin exposures have been assessed. However, such a study would cost 120,000 € when including 120 participants.

Effects of dioxin exposure on breast cancer risk may depend on whether the exposure occurred during critical periods of mammary development (*in utero*, puberty, pregnancy and lactation) (Jenkins et al. 2012; Teitelbaum et al. 2015). Because E3N women were aged 45 to 60 years at baseline, and given the lack of dioxin measurement data before 1990, we could not assess both dietary and airborne dioxin exposures at these time-periods, thus our study may have missed critical windows. However, information regarding age at menarche, age at first-full term pregnancy and breastfeeding of the E3N women were available and we were able to estimate breast cancer risk associated with dioxin exposure among subgroups defined by these factors.

Existing literature on dioxin exposure and breast cancer only focused on airborne dioxin exposure; however, the main route of dioxin exposure is through ingestion of contaminated food. In our research project, we assessed both dietary and airborne dioxin and estimated breast cancer risk separately for

the two routes. Although women were exposed to both dietary and airborne dioxin exposures at once, we did not investigate multiple-pathway dioxin exposure, and could not observe the potential joint effect of multiple dioxin exposure and breast cancer risk. However, logistic regression models estimating breast cancer risk associated with airborne dioxin exposure were further adjusted for dietary dioxin exposure, and this did not change our risk estimates. Statistical models have been developed to consider multiple pathways of exposure in risk estimation (Petit et al. 2012); future research should consider the joint effects of multiple-pathway exposures on health outcomes.

Experimental testing of the combined effects of a mixture of xenoestrogens showed that they were able to act together and produce significant effects even when individual components were present at low concentration, suggesting the importance of considering the joint action of estrogenic chemicals to avoid potential underestimation of risk estimates (Silva et al. 2002). These observations can be transposed to humans, which are exposed to complex and changing mixtures of environmental compounds at once. These compounds may influence independently, cumulatively or interactively occurrence of cancer. Recent reviews concluded that no universal multipollutant exposure metric existed regarding exposure assessment and that specific statistical models should be developed and applied in order to consider multicollinearity, interaction and the large number of pollutants when estimating the association between exposure and health outcome (Billionnet et al. 2012; Oakes et al. 2014). Industrial facilities emit not only dioxins but other air pollutants that can act as EDCs, such as cadmium. Our study focused on dioxins and may have missed this complex effect of multipollutant exposure on breast cancer risk. Although no clear consensus has yet emerged from the literature on how a multipollutant approach should be addressed in epidemiological studies, future research should be directed toward this scope.

2. Research perspectives

2.1. The GEO3N research project

The doctoral research conducted in the GEO3N project among women permanently residing in the Rhône-Alpes region from 1990 to 2008 allowed the development of a GIS-based metric for the assessment of cumulative airborne dioxin exposure and the implementation of specific statistical models to estimate time-dependent breast cancer risks in a nested case-control study. The study is currently being extended to the whole French territory, among nearly 5,500 breast cancer cases and 5,500 matched controls. The retrospective inventory and characterization of the industrial sources was achieved recently; as of today, the GIS-based analysis remains to be performed in order to estimate the association between airborne dioxin exposure and breast cancer risk. Moreover, the questionnaire on lifetime residential history has been modified and will be mailed to the women (not deceased) of the GEO3N study population (except for women of the Rhône-Alpes region) by the end of 2016. Two reminders have been planned for the non-responding women.

The Institute of Medicine recently advised the development of a life course approach to explore the long-term effects of environmental exposures at different time-points in the lifetime on women's risk of breast cancer (IOM (Institute of Medicine) 2012). In our study on dioxin exposure and breast cancer risk, the estimated dioxin intake was assessed at one time-point in 1993, and airborne dioxin exposure was assessed for each year between 1990 and 2008. Regarding dioxins in the French context, no data on food contamination by dioxins were available before 1990, and information regarding historical emission of dioxins from industrial sources is lacking as well; thus, we could not study the effects of dioxin exposure over women's lifetime. However, a re-assessment of dietary dioxin exposure within the E3N cohort will be conceivable in future studies, using food consumption data from the second E3N diet history questionnaire (not available at the time of analysis) and food

contamination data from the TDS (Sirot et al. 2009; Tard et al. 2007), and breast cancer risk will be estimated regarding the evolution in dietary dioxin exposure of the study population. In a previous study investigating life in an urban area and breast cancer risk in the E3N cohort, being born in an urban area was associated with increased breast cancer risk (Binachon et al. 2014) (Appendice 6). Other studies suggested the use of residence in urban places as a surrogate for air pollution exposure released from industrial facilities or traffic (Lambrechtsen et al. 2012). Moreover, the lifetime residential history of the women included in the nested case-control study has been recently collected through the QHR (Appendice 4) and assigned an “urban” or “rural” status; the mailing of QHR will be extended to the women of the whole French territory. Therefore, lifetime airborne dioxin exposure could be assessed using urban residence as a surrogate for earlier periods when airborne dioxin measurements were not available.

The association between dioxin exposure and breast cancer risk may be modified by individual differences in genetic susceptibility of women. In the field of cancer, research has focused on genetic polymorphisms of phase I and phase II enzymes, which are involved in the metabolism of xenobiotics, as well as in the hormone metabolism. Epidemiological studies published so far on gene-environment interactions between genetic polymorphisms and dioxin exposure have focused on enzymes from the CYP1 detoxication system (CYP1A1, CYP1A2, CYP1B1, CYP2S1). A French study reported a statistically significant interaction between a SNP of the CYP1B1 gene, associated with change in catalytic function, and dioxin exposure in a case-series of 283 breast cancer patients. Women carrying the variant allele and who had lived in the vicinity of a MSWI for more than ten years had a higher risk of breast cancer than women never exposed and homozygous for the major allele (OR for interaction=3.3, 95%CI=1.2-8.8) (Saintot et al. 2004). A study conducted among 121 subjects exposed to TCDD after the Seveso accident found no effect modification of CYP1A1 and CYP1B1 variants on the association between TCDD plasma and gene expression. Additional *in vitro* analysis of lymphocytes treated with TCDD showed that gene expression of CYP1A1 and CYP1B1 was modified by variants alleles (Landi et al. 2005). Recently in a study of 421 healthy pregnant Japanese women, statistically significant differences, although relatively weak, in dioxin blood concentrations were reported for polymorphisms in AhR and CYP1A1 (Kobayashi et al. 2013). Other studies focused on gene-environment interactions and breast cancer risk; however, none examined interactions between genetic polymorphisms and dioxin exposure (Barrdahl et al. 2014; Nickels et al. 2013). Inserm and Anses advised the further research on gene-environment interactions, with the development of studies using quality data on environmental exposure and a large number of SNPs available among both cases and controls (Inserm and Afsset 2008). The aspect of gene-environment interaction is being implemented within the GEO3N study population, among 2,500 breast cancer cases and 2,500 matched controls for whom DNA was available. Research will focus on genetic polymorphisms involved in dioxin metabolic pathways, as well as polymorphisms of genes encoding the estrogen receptor and growth factors to explore gene-environment interactions. Genotyping will be achieved using the HumanCore Exom DNA microarray (Centre National de Génotypage, Evry, France) which allows covering the entire genome.

2.2. The XENAIR research project

The Cancer and Environment research department and the team “Lifestyle, genes and health” have extended their collaboration and developed the XENAIR research project that aims at investigating long-term effects of exposure to multiple ambient air pollutants with xenoestrogen properties (particulate matter, nitrogen dioxide, ozone, B[a]P, dioxins, PCBs and cadmium) and breast cancer risk in a case-control study nested within the French E3N prospective cohort. The project proposes to address the hypotheses that (i) breast cancer risk is associated with chronic exposure to ambient air

pollutants; (ii) this association may be stronger with increasing years of exposure and among susceptible subgroups; and (iii) the association may vary according to timing of exposure and hormone receptor status.

The specific objectives concern the assessment of environmental exposure at several spatial resolutions, using GIS, LUR and dispersion modeling; to investigate how long-term environmental exposure may impact breast cancer risk, by estimating the relative weight of exposure with respect to time and the cumulative duration of urban residence since birth, assessing indoor exposure from domestic heating and combustion activities, and exploring approaches to estimate multi-pollutant exposure. Interactions between long-term exposure to ambient air pollutants and genetic polymorphisms involved in the metabolism of xenoestrogens will be tested among 2,500 breast cancer cases and 2,500 matched controls. The role of DNA methylation as an intermediate marker of risk in the association between exposure to ambient air pollutants and breast cancer risk will be examined in a nested case-control study including 200 non-smoking breast cancer cases matched to 200 non-smoking controls that provided DNA. The fraction of breast cancer attributable to air pollution in France will be analyzed based on the risk estimates and the costs associated with the management of breast cancer attributable to air pollutants will be quantified.

The XENAIR research project received a grant from the Fondation ARC pour la recherche sur le cancer (CANC’AIR), and the GIS-based metric is currently being adapted for the assessment of cadmium exposure.

3. Public Health impact

The increase in breast cancer incidence observed in France and worldwide, has become a major Public Health concern. Research on breast cancer risk factors is still ongoing and is of great importance, as it has allowed the identification and implementation of preventive actions that women can adopt to potentially reduce their risk of breast cancer: organized screening; minimizing medical radiation; avoiding use of postmenopausal hormone therapy combining estrogen and progesterone; avoiding smoking; limiting alcohol consumption; increasing physical activity and minimizing weight gain among others (IOM (Institute of Medicine) 2012).

In France, the relationship between cancer and environment has become a Public Health concern and a growing query from the population and health professionals, particularly because environmental exposures are common and avoidable. Several national, regional and local plans have been implemented that aim at supporting research and prevention of cancers related to environmental factors at the national level (INCa 2009, 2014), and preserving health, improving environmental quality and advising the public at the regional level (PNSE2 2009). On the local level, the Plan regional santé-environnement of the Rhône-Alpes region aims at reducing exposure to emissions from human activities, controlling nuisances related to transport, managing the industrial footprint and reducing risks associated with indoor environments, among others (PRSE2 2011). Moreover, since 2006, Air Rhône-Alpes, a regional association on air quality monitoring, is conducting a program for monitoring of dioxins and heavy metals to inform the public about real time pollution episodes (Air Rhône-Alpes 2016).

As a response to the recommendations of these plans, the Cancer and Environment research department has implemented a web portal in order to inform patients, health professionals and the public about the relationship between health and environment (Cancer Environnement). This website aims at facilitating access to scientific, medical and technical information; developing knowledge and skills of different audiences; developing cooperation between the different actors in Environment, Oncology, Research and Prevention. Several factsheets have been elaborated, by cancer sites and carcinogens (occupational and environmental); including two regarding dioxins, explaining their

origin, their exposure sources and their known effects on cancers and other pathologies.

The research program of the Cancer and Environment research department also focuses on the investigation of environmental risk factors associated with breast cancer and in particular on the impact of waste management. In this context, and with the collaboration of the team “Lifestyle, genes and health”, two major research projects are currently ongoing, through an interdisciplinary research, in order to clarify the relationship between environmental pollutants with endocrine-disrupting properties and breast cancer risk. Preventive actions will be considered if positive associations between air pollutant exposure and breast cancer risk are observed.

4. Conclusion

Research has recently focused on the environmental pollutants as potential risk factors to explain the increase in breast cancer incidence observed in France and worldwide. However, findings remain inconclusive due to differences in methodology across studies and study limitations. This doctoral research intended to clarify the relationship between dioxin exposure and breast cancer risk in the general population by considering the combined exposure of individuals to dioxins *via* multiple and relevant routes, providing new knowledge through an interdisciplinary research and filling the methodological gaps of the existing scientific literature. Moreover, research conducted in the GEO3N project allowed the development of a standardized GIS-based tool for the assessment of airborne dioxin exposure. This tool will be adapted for the exposure assessment of other environmental pollutants in relation to breast cancer or applied in the investigation of other pathologies.

References

- Air Rhône-Alpes. 2016. La qualité de l'air en Auvergne-Rhône-Alpes. Air Rhône-Alpes. Available: <http://www.air-rhonealpes.fr/> [accessed 9 October 2016].
- Alexander DD, Morimoto LM, Mink PJ, Cushing CA. 2010. A review and meta-analysis of red and processed meat consumption and breast cancer. *Nutr. Res. Rev.* 23:349–365; doi:10.1017/S0954422410000235.
- American Cancer Society. 2015. Breast Cancer Facts & Figures 2015-2016.
- American Institute for Cancer Research, World Cancer Research Fund, eds. 2010. *Continuous Update Project Report: Food, Nutrition, Physical Activity and the Prevention of Breast Cancer*.
- Anderson WF, Rosenberg PS, Prat A, Perou CM, Sherman ME. 2014. How Many Etiological Subtypes of Breast Cancer: Two, Three, Four, Or More? *J. Natl. Cancer Inst.* 106:dju165; doi:10.1093/jnci/dju165.
- Antoniou A, Pharoah PDP, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. 2003. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am. J. Hum. Genet.* 72: 1117–1130.
- Anzivino-Viricel L, Falette N, Carretier J, Montestrucq L, Guye O, Philip T, et al. 2012. Gestion des déchets ménagers et assimilés: bilan des connaissances et évaluation des effets sanitaires en population générale et au travail. *Environ. Risques Santé* 11: 360–377.
- Apostolou P, Fostira F. 2013. Hereditary Breast Cancer: The Era of New Susceptibility Genes. *BioMed Res. Int.* 2013; doi:10.1155/2013/747318.
- Arisawa K, Matsumura T, Tohyama C, Saito H, Satoh H, Nagai M, et al. 2003. Fish intake, plasma ω -3 polyunsaturated fatty acids, and polychlorinated dibenzo-p-dioxins/polychlorinated dibenzo-furans and co-planar polychlorinated biphenyls in the blood of the Japanese population. *Int. Arch. Occup. Environ. Health* 76: 205–215.
- Arisawa K, Uemura H, Hiyoshi M, Kitayama A, Takami H, Sawachika F, et al. 2011. Dietary patterns and blood levels of PCDDs, PCDFs, and dioxin-like PCBs in 1656 Japanese individuals. *Chemosphere* 82:656–662; doi:10.1016/j.chemosphere.2010.11.008.
- Aronson KJ, Miller AB, Woolcott CG, Sterns EE, McCready DR, Lickley LA, et al. 2000. Breast adipose tissue concentrations of polychlorinated biphenyls and other organochlorines and breast cancer risk. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 9: 55–63.
- Ashida H, Fukuda I, Yamashita T, Kanazawa K. 2000. Flavones and flavonols at dietary levels inhibit a transformation of aryl hydrocarbon receptor induced by dioxin. *FEBS Lett.* 476: 213–217.
- Aune D, Chan DSM, Greenwood DC, Vieira AR, Rosenblatt DAN, Vieira R, et al. 2012a. Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Ann. Oncol.* 23:1394–1402; doi:10.1093/annonc/mdr589.
- Aune D, Chan DSM, Vieira AR, Rosenblatt DAN, Vieira R, Greenwood DC, et al. 2012b. Fruits, vegetables and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Breast Cancer Res. Treat.* 134:479–493; doi:10.1007/s10549-012-2118-1.
- Baan R, Grosse Y, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. 2009. A review of human carcinogens--Part F: chemical agents and related occupations. *Lancet Oncol.* 10: 1143–1144.
- Barrdahl M, Canzian F, Joshi AD, Travis RC, Chang-Claude J, Auer PL, et al. 2014. Post-GWAS gene-environment interplay in breast cancer: results from the Breast and Prostate Cancer Cohort Consortium and a meta-analysis on 79 000 women. *Hum. Mol. Genet.* 23:5260–5270; doi:10.1093/hmg/ddu223.

- Basagaña X, Aguilera I, Rivera M, Agis D, Foraster M, Marrugat J, et al. 2013. Measurement Error in Epidemiologic Studies of Air Pollution Based on Land-Use Regression Models. *Am. J. Epidemiol.* 178:1342–1346; doi:10.1093/aje/kwt127.
- Beckmann MW, Niederacher D, Schnürch H-G, Gusterson BA, Bender HG. 1997. Multistep carcinogenesis of breast cancer and tumour heterogeneity. *J. Mol. Med.* 75: 429–439.
- Belot A, Grosclaude P, Bossard N, Jouglu E, Benhamou E, Delafosse P, et al. 2008. Cancer incidence and mortality in France over the period 1980–2005. *Rev. Dépidémiologie Santé Publique* 56: 159–175.
- Benz CC. 2008. Impact of aging on the biology of breast cancer. *Crit. Rev. Oncol. Hematol.* 66:65–74; doi:10.1016/j.critrevonc.2007.09.001.
- Bernstein L. 2002. Epidemiology of endocrine-related risk factors for breast cancer. *J. Mammary Gland Biol. Neoplasia* 7: 3–15.
- Bertazzi A, Pesatori AC, Consonni D, Tironi A, Landi MT, Zocchetti C. 1993. Cancer incidence in a population accidentally exposed to 2,3,7,8-tetrachlorodibenzo-para-dioxin. *Epidemiol. Camb. Mass* 4: 398–406.
- Bilau M, Matthys C, Baeyens W, Bruckers L, Backer GD, Hond ED, et al. 2008. Dietary exposure to dioxin-like compounds in three age groups: Results from the Flemish environment and health study. *Chemosphere* 70:584–592; doi:10.1016/j.chemosphere.2007.07.008.
- Billionnet C, Sherrill D, Annesi-Maesano I, GERIE study. 2012. Estimating the health effects of exposure to multi-pollutant mixture. *Ann. Epidemiol.* 22:126–141; doi:10.1016/j.annepidem.2011.11.004.
- Binachon B, Dossus L, Danjou AMN, Clavel-Chapelon F, Fervers B. 2014. Life in urban areas and breast cancer risk in the French E3N cohort. *Eur. J. Epidemiol.* 29:743–751; doi:10.1007/s10654-014-9942-z.
- Bodian CA. 1993. Benign breast diseases, carcinoma in situ, and breast cancer risk. *Epidemiol. Rev.* 15: 177–187.
- Boyd NF, Stone J, Vogt KN, Connelly BS, Martin LJ, Minkin S. 2003. Dietary fat and breast cancer risk revisited: a meta-analysis of the published literature. *Br. J. Cancer* 89:1672–1685; doi:10.1038/sj.bjc.6601314.
- Brennan SF, Cantwell MM, Cardwell CR, Velentzis LS, Woodside JV. 2010. Dietary patterns and breast cancer risk: a systematic review and meta-analysis. *Am. J. Clin. Nutr.* 91:1294–1302; doi:10.3945/ajcn.2009.28796.
- Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. 2007. Environmental pollutants and breast cancer: Epidemiologic studies. *Cancer* 109:2667–2711; doi:10.1002/cncr.22655.
- Brody JG, Rudel RA. 2003. Environmental Pollutants and Breast Cancer. *Environ. Health Perspect.* 111:1007–1019; doi:10.1289/ehp.6310.
- Brunnberg S, Andersson P, Poellinger L, Hanberg A. 2011. The constitutively active Ah receptor (CA-AhR) mouse as a model for dioxin exposure – Effects in reproductive organs. *Chemosphere* 85:1701–1706; doi:10.1016/j.chemosphere.2011.09.015.
- Cancer Environnement. Portail officiel des risques de cancer (professionnels, alimentaires, environnementaux) cancer lyon | Cancer et environnement. Available: <http://www.cancer-environnement.fr/> [accessed 10 October 2016].
- Centre Léon Bérard. 2016. Centre Léon Bérard | Centre de lutte contre le cancer Lyon, Rhône-Alpes, traitement et soins cancer, recherche cancérologie. Available: <http://www.centreleonberard.fr/> [accessed 29 September 2016].
- Chen F-P, Chien M-H. 2014. Lower concentrations of phthalates induce proliferation in human breast cancer cells. *Climacteric J. Int. Menopause Soc.* 17:377–384; doi:10.3109/13697137.2013.865720.

- Chen HL, Su HJ, Lee CC. 2007. Association between tofu intake and serum polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs) in the elderly Taiwanese. *Environ. Int.* 33:265–271; doi:10.1016/j.envint.2006.09.019.
- Cho E, Chen WY, Hunter DJ, et al. 2006. Red meat intake and risk of breast cancer among premenopausal women. *Arch. Intern. Med.* 166:2253–2259; doi:10.1001/archinte.166.20.2253.
- Chowdhury R, Sinha B, Sankar MJ, Taneja S, Bhandari N, Rollins N, et al. 2015. Breastfeeding and maternal health outcomes: a systematic review and meta-analysis. *Acta Paediatr. Oslo Nor.* 1992 104:96–113; doi:10.1111/apa.13102.
- CITEPA. 2015. Dioxines et Furanes - PCDD-F. Available: <http://www.citepa.org/fr/air-et-climat/polluants/polluant-organiques-persistants/dioxines-et-furanes> [accessed 1 September 2016].
- Clavel-Chapelon F, for the E3N Study Group. 2015. Cohort Profile: The French E3N Cohort Study. *Int. J. Epidemiol.* 44:801–809; doi:10.1093/ije/dyu184.
- Clavel-Chapelon F, Gerber M. 2002. Reproductive factors and breast cancer risk. Do they differ according to age at diagnosis? *Breast Cancer Res. Treat.* 72: 107–115.
- Clavel-Chapelon F, van Liere MJ, Giubout C, Niravong MY, Goulard H, Le Corre C, et al. 1997. E3N, a French cohort study on cancer risk factors. E3N Group. *Etude Epidémiologique auprès de femmes de l'Education Nationale.* *Eur. J. Cancer Prev. Off. J. Eur. Cancer Prev. Organ. ECP* 6: 473–478.
- Collaborative Group on Hormonal Factors in Breast Cancer. 1997. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52 705 women with breast cancer and 108 411 women without breast cancer. *The Lancet* 350: 1047–1059.
- Collaborative Group on Hormonal Factors in Breast Cancer, others. 2002. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96 973 women without the disease. *The Lancet* 360: 187–195.
- Collaborative Group on Hormonal Factors in Breast Cancer, others. 1996. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. *The Lancet* 347: 1713–1727.
- Collaborative Group on Hormonal Factors in Breast Cancer, others. 2001. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58 209 women with breast cancer and 101 986 women without the disease. *The Lancet* 358: 1389–1399.
- Collaborative Group on Hormone Factors and Breast Cancer. 2002. Alcohol, tobacco and breast cancer – collaborative reanalysis of individual data from 53 epidemiological studies, including 58 515 women with breast cancer and 95 067 women without the disease. *Br. J. Cancer* 87:1234–1245; doi:10.1038/sj.bjc.6600596.
- Collaborators on the Million Women Study. 2003. Breast cancer and hormone-replacement therapy in the Million Women Study. *The Lancet* 362:419–427; doi:10.1016/S0140-6736(03)14065-2.
- Consonni D, Sindaco R, Bertazzi PA. 2012. Blood levels of dioxins, furans, dioxin-like PCBs, and TEQs in general populations: A review, 1989–2010. *Environ. Int.* 44:151–162; doi:10.1016/j.envint.2012.01.004.
- Copeland KT, Checkoway H, McMICHAEL AJ, Holbrook RH. 1977. Bias Due to Misclassification in the Estimation of Relative Risk. *Am. J. Epidemiol.* 105: 488–495.
- Cordioli M, Ranzi A, De Leo GA, Lauriola P. 2013. A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators. *J. Environ. Public Health* 2013; doi:10.1155/2013/129470.
- Coronado GD, Beasley J, Livaudais J. 2011. Alcohol consumption and the risk of breast cancer. *Salud Pública México* 53: 440–447.

- Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F, et al. 2009. Postmenopausal Breast Cancer Risk and Dietary Patterns in the E3N-EPIC Prospective Cohort Study. *Am. J. Epidemiol.* 170:1257–1267; doi:10.1093/aje/kwp257.
- CSHPF. 2000. Dioxines: données de contamination et d'exposition de la population française.
- Cui Y, Miller AB, Rohan TE. 2006. Cigarette smoking and breast cancer risk: update of a prospective cohort study. *Breast Cancer Res. Treat.* 100:293–299; doi:10.1007/s10549-006-9255-3.
- Dai D, Oyana TJ. 2008. Spatial variations in the incidence of breast cancer and potential risks associated with soil dioxin contamination in Midland, Saginaw, and Bay Counties, Michigan, USA. *Environ. Health* 7:49; doi:10.1186/1476-069X-7-49.
- Danjou A, Fervers B, Boutron-Ruault M-C, Philip T, Clavel-Chapelon F, Dossus L. 2015. Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort. *Breast Cancer Res.* 17:39; doi:10.1186/s13058-015-0536-9.
- Dartois L, Fagherazzi G, Baglietto L, Boutron-Ruault M-C, Delaloue S, Mesrine S, et al. 2016. Proportion of premenopausal and postmenopausal breast cancers attributable to known risk factors: Estimates from the E3N-EPIC cohort. *Int. J. Cancer* 138:2415–2427; doi:10.1002/ijc.29987.
- de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, et al. 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environ. Int.* 73:382–392; doi:10.1016/j.envint.2014.08.011.
- De Mul A, Bakker MI, Zeilmaker MJ, Traag WA, Leeuwen SPJ van, Hoogenboom RLAP, et al. 2008. Dietary exposure to dioxins and dioxin-like PCBs in The Netherlands anno 2004. *Regul. Toxicol. Pharmacol.* 51:278–287; doi:10.1016/j.yrtph.2008.04.010.
- Demers A, Ayotte P, Brisson J, Dodin S, Robert J, Dewailly É. 2002. Plasma Concentrations of Polychlorinated Biphenyls and the Risk of Breast Cancer: A Congener-specific Analysis. *Am. J. Epidemiol.* 155:629–635; doi:10.1093/aje/155.7.629.
- Den Hond E, Roels HA, Hoppenbrouwers K, Nawrot T, Thijs L, Vandermeulen C, et al. 2002. Sexual maturation in relation to polychlorinated aromatic hydrocarbons: Sharpe and Skakkebaek's hypothesis revisited. *Environ. Health Perspect.* 110: 771–776.
- di Domenico A, Cerlesi S, Ratti S. 1990. A two-exponential model to describe the vanishing trend of 2,3,7,8-tetrachlorodibenzodioxin (TCDD) in the soil at Seveso, northern Italy. *Chemosphere* 20:1559–1566; doi:10.1016/0045-6535(90)90313-I.
- Diamanti-Kandarakis E, Bourguignon J-P, Giudice LC, Hauser R, Prins GS, Soto AM, et al. 2009. Endocrine-Disrupting Chemicals: An Endocrine Society Scientific Statement. *Endocr. Rev.* 30:293–342; doi:10.1210/er.2009-0002.
- Domingo JL, Schuhmacher M, Granero S, Llobet JM. 1999. PCDDs and PCDFs in food samples from Catalonia, Spain. An assessment of dietary intake. *Chemosphere* 38: 3517–3528.
- Dong J-Y, Zhang L, He K, Qin L-Q. 2011. Dairy consumption and risk of breast cancer: a meta-analysis of prospective cohort studies. *Breast Cancer Res. Treat.* 127:23–31; doi:10.1007/s10549-011-1467-5.
- Dossus L, Boutron-Ruault M-C, Kaaks R, Gram IT, Vilier A, Fervers B, et al. 2014. Active and passive cigarette smoking and breast cancer risk: results from the EPIC cohort. *Int. J. Cancer* 134:1871–1888; doi:10.1002/ijc.28508.
- Dyrstad SW, Yan Y, Fowler AM, Colditz GA. 2015. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. *Breast Cancer Res. Treat.* 149:569–575; doi:10.1007/s10549-014-3254-6.
- Education Nationale. Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche. Available: <http://www.education.gouv.fr/> [accessed 10 October 2016].

- El Ghissassi F, Baan R, Straif K, Grosse Y, Secretan B, Bouvard V, et al. 2009. A review of human carcinogens—Part D: radiation. *Lancet Oncol.* 10:751–752; doi:10.1016/S1470-2045(09)70213-X.
- Fabre P, Daniau C, Gorla S, de Crouy-Chanel P, Empereur-Bissonnet P. 2008. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. Synthèse St.-Maurice InVS.
- Fachal L, Dunning AM. 2015. From candidate gene studies to GWAS and post-GWAS analyses in breast cancer. *Curr. Opin. Genet. Dev.* 30:32–41; doi:10.1016/j.gde.2015.01.004.
- FAO/WHO. 2005. Dietary exposure assessment of chemicals in food. Report of a joint FAO/WHO consultation.
- Fenton SE, Birnbaum LS. 2015. Timing of Environmental Exposures as a Critical Element in Breast Cancer Risk. *J. Clin. Endocrinol. Metab.* 100:3245–3250; doi:10.1210/jc.2015-2848.
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. 2015. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136:E359–E386; doi:10.1002/ijc.29210.
- Feychting M, Forssén U. 2006. Electromagnetic Fields and Female Breast Cancer. *Cancer Causes Control* 17:553–558; doi:10.1007/s10552-005-9008-3.
- Fontham ETH, Thun MJ, Ward E, Balch AJ, Delancey JOL, Samet JM, et al. 2009. American Cancer Society Perspectives on Environmental Factors and Cancer. *CA. Cancer J. Clin.* 59:343–351; doi:10.3322/caac.20041.
- Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJ, et al. 2011. Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *The lancet* 378: 1461–1484.
- Fournier A, Mesrine S, Dossus L, Boutron-Ruault M-C, Clavel-Chapelon F, Chabbert-Buffet N. 2014. Risk of breast cancer after stopping menopausal hormone therapy in the E3N cohort. *Breast Cancer Res. Treat.* 145:535–543; doi:10.1007/s10549-014-2934-6.
- Fréry N, Zeghnoun A, Sarter H, Falq G, Pascal M, Bérat B, et al. 2009. *Etude d'imprégnation par les dioxines des populations vivant à proximité d'usines d'incinération d'ordures ménagères - Rapport d'étude.* Institut de veille sanitaire. Saint-Maurine (Fra).
- Friedenreich CM. 2001. Review of anthropometric factors and breast cancer risk. *Eur. J. Cancer Prev.* 10: 15–32.
- Fries GF. 1995. A review of the significance of animal food products as potential pathways of human exposures to dioxins. *J. Anim. Sci.* 73: 1639–1650.
- Gammon MD, Santella RM, Neugut AI, Eng SM, Teitelbaum SL, Paykin A, et al. 2002a. Environmental toxins and breast cancer on Long Island. I. Polycyclic aromatic hydrocarbon DNA adducts. *Cancer Epidemiol. Biomarkers Prev.* 11: 677–685.
- Gammon MD, Wolff MS, Neugut AI, Eng SM, Teitelbaum SL, Britton JA, et al. 2002b. Environmental toxins and breast cancer on Long Island. II. Organochlorine compound levels in blood. *Cancer Epidemiol. Biomarkers Prev.* 11: 686–697.
- Garcia E, Hurley S, Nelson DO, Hertz A, Reynolds P. 2015. Hazardous air pollutants and breast cancer risk in California teachers: a cohort study. *Environ. Health* 14:14; doi:10.1186/1476-069X-14-14.
- Giulivo M, Lopez de Alda M, Capri E, Barceló D. 2016. Human exposure to endocrine disrupting compounds: Their role in reproductive systems, metabolic syndrome and breast cancer. A review. *Environ. Res.* 151:251–264; doi:10.1016/j.envres.2016.07.011.
- Grassman JA, Masten SA, Walker NJ, Lucier GW. 1998. Animal models of human response to dioxins. *Environ. Health Perspect.* 106: 761–775.

- Greenberg ER, Barnes AB, Resseguie L, Barrett JA, Burnside S, Lanza LL, et al. 1984. Breast Cancer in Mothers Given Diethylstilbestrol in Pregnancy. *N. Engl. J. Med.* 311:1393–1398; doi:10.1056/NEJM198411293112201.
- Gulliver J, Briggs D. 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. *Sci. Total Environ.* 409:2419–2429; doi:10.1016/j.scitotenv.2011.03.004.
- Guo S-W, Simsa P, Kyama CM, Mihalyi A, Fulop V, Othman E-ER, et al. 2009. Reassessing the evidence for the link between dioxin and endometriosis: from molecular biology to clinical epidemiology. *Mol. Hum. Reprod.* 15:609–624; doi:10.1093/molehr/gap075.
- Hagmar L, Mikoczy Z, Welinder H. 1995. Cancer incidence in Swedish sterilant workers exposed to ethylene oxide. *Occup. Environ. Med.* 52: 154–156.
- Han D, Bonner MR, Nie J, Freudenheim JL. 2013. Assessing bias associated with geocoding of historical residence in epidemiology research. *Geospatial Health* 7:369–374; doi:10.4081/gh.2013.93.
- Hannaford PC, Selvaraj S, Elliott AM, Angus V, Iversen L, Lee AJ. 2007. Cancer risk among users of oral contraceptives: cohort data from the Royal College of General Practitioner’s oral contraception study. *BMJ* 335:651–651; doi:10.1136/bmj.39289.649410.55.
- Hardell L, Lindström G, Liljegren G, Dahl P, Magnuson A. 1996. Increased concentrations of octachlorodibenzo-p-dioxin in cases with breast cancer--results from a case-control study. *Eur. J. Cancer Prev. Off. J. Eur. Cancer Prev. Organ. ECP* 5: 351–357.
- Hattemer-Frey HA, Travis CC. 1989. Chlorinated Dioxins and Related Compounds Comparison of human exposure to dioxin from municipal waste incineration and background environmental contamination. *Chemosphere* 18:643–649; doi:10.1016/0045-6535(89)90177-X.
- Hauptmann M, Wellmann J, Lubin JH, Rosenberg PS, Kreienbrock L. 2000. Analysis of Exposure-Time-Response Relationships Using a Spline Weight Function. *Biometrics* 56: 1105–1108.
- He C, Anand ST, Ebell MH, Vena JE, Robb SW. 2014. Circadian disrupting exposures and breast cancer risk: a meta-analysis. *Int. Arch. Occup. Environ. Health* 88:533–547; doi:10.1007/s00420-014-0986-x.
- Hoek G, Fischer P, Van Den Brandt P, Goldbohm S, Brunekreef B. 2001. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J. Expo. Anal. Environ. Epidemiol.* 11.
- Hoover RN, Hyer M, Pfeiffer RM, Adam E, Bond B, Cheville AL, et al. 2011. Adverse Health Outcomes in Women Exposed In Utero to Diethylstilbestrol. *N. Engl. J. Med.* 365:1304–1314; doi:10.1056/NEJMoa1013961.
- Hsieh T-H, Tsai C-F, Hsu C-Y, Kuo P-L, Lee J-N, Chai C-Y, et al. 2012. Phthalates induce proliferation and invasiveness of estrogen receptor-negative breast cancer through the AhR/HDAC6/c-Myc signaling pathway. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 26:778–787; doi:10.1096/fj.11-191742.
- Hung M-C, Hwang J-J. 2013. Cancer risk from medical radiation procedures for coronary artery disease: a nationwide population-based cohort study. *Asian Pac. J. Cancer Prev. APJCP* 14: 2783–2787.
- IARC, ed. 2012a. *IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 D, radiation: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 02 - 09 June 2009.* IARC, Lyon.
- IARC, ed. 2012b. *IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 E, personal habits and indoor combustions: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 29 September - 06 October 2009.* IARC, Lyon.

- IARC, ed. 2012c. *IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 F, chemical agents and related occupations: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 20 - 27 October 2009*. IARC, Lyon.
- IARC, ed. 2016. *IARC monographs on the evaluation of carcinogenic risks to humans, volume 107, polychlorinated biphenyls and polybrominated biphenyls: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 12 - 19 February 2013*. IARC, Lyon.
- IARC, ed. 1997. *Polychlorinated dibenzo-para-dioxins and polychlorinated dibenzofurans: ... views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 4 - 11 February 1997*. Lyon.
- INCa. 2009. Le Plan cancer 2009-2013 - Les Plans cancer de 2003 à 2013 | Institut National Du Cancer. Available: <http://www.e-cancer.fr/Plan-cancer/Les-Plans-cancer-de-2003-a-2013/Le-Plan-cancer-2009-2013> [accessed 4 October 2016].
- INCa. 2014. Plan cancer 2014-2019 : priorités et objectifs - Plan cancer | Institut National Du Cancer. Available: <http://www.e-cancer.fr/Plan-cancer/Plan-cancer-2014-2019-priorites-et-objectifs> [accessed 3 October 2016].
- INCa. 2015. Projection de l'incidence et de la mortalité par cancer en France métropolitaine en 2015. Rapport technique.
- Insee. 2016. Definitions and methods - Urban unit. Available: <http://www.insee.fr/en/methodes/default.asp?page=definitions/unite-urbaine.htm> [accessed 2 October 2016].
- Inserm, Afsset. 2008. *Cancer et environnement - Expertise collective*. Les éditions Inserm.
- Inserm U1018. 2016. Cohorte E3N. Available: <http://e3n.fr/> [accessed 29 September 2016].
- IOM (Institute of Medicine). 2012. *Breast Cancer and the Environment: A Life Course Approach*. National Academies Press, Washington, D.C.
- Jenkins S, Betancourt AM, Wang J, Lamartiniere CA. 2012. Endocrine-active chemicals in mammary cancer causation and prevention. *J. Steroid Biochem. Mol. Biol.* 129:191–200; doi:10.1016/j.jsbmb.2011.06.003.
- John EM, McGuire V, Thomas D, Haile R, Ozcelik H, Milne RL, et al. 2013. Diagnostic chest X-rays and breast cancer risk before age 50 years for BRCA1 and BRCA2 mutation carriers. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 22:1547–1556; doi:10.1158/1055-9965.EPI-13-0189.
- Johnson KC. 2005. Accumulating evidence on passive and active smoking and breast cancer risk. *Int. J. Cancer* 117:619–628; doi:10.1002/ijc.21150.
- JORF n°280. Arrêté du 20 septembre 2002 relatif aux installations d'incinération et de co-incinération de déchets non dangereux et aux installations incinérant des déchets d'activités de soins à risques infectieux. NOR: DEVP0210351A: 19778.
- Jung S, Spiegelman D, Baglietto L, Bernstein L, Boggs DA, Brandt PA van den, et al. 2013. Fruit and Vegetable Intake and Risk of Breast Cancer by Hormone Receptor Status. *J. Natl. Cancer Inst.* 105:219–236; doi:10.1093/jnci/djs635.
- Kamrin M. 2007. The “Low Dose” Hypothesis: Validity and Implications for Human Risk. *Int. J. Toxicol.* 26:13–23; doi:10.1080/10915810601117968.
- Kavlock RJ, Daston GP, DeRosa C, Fenner-Crisp P, Gray LE, Kaattari S, et al. 1996. Research needs for the risk assessment of health and environmental effects of endocrine disruptors: a report of the U.S. EPA-sponsored workshop. *Environ. Health Perspect.* 104: 715–740.

- Keen JC, Davidson NE. 2003. The biology of breast carcinoma. *Cancer* 97:825–833; doi:10.1002/cncr.11126.
- Kliwer EV, Smith KR. 1995. Breast cancer mortality among immigrants in Australia and Canada. *J. Natl. Cancer Inst.* 87: 1154–1161.
- Kobayashi S, Sata F, Sasaki S, Ban S, Miyashita C, Okada E, et al. 2013. Genetic association of aromatic hydrocarbon receptor (AHR) and cytochrome P450, family 1, subfamily A, polypeptide 1 (CYP1A1) polymorphisms with dioxin blood concentrations among pregnant Japanese women. *Toxicol. Lett.* 219:269–278; doi:10.1016/j.toxlet.2013.03.013.
- Kociba RJ, Keyes DG, Beyer JE, Carreon RM, Wade CE, Dittenber DA, et al. 1978. Results of a two-year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in rats. *Toxicol. Appl. Pharmacol.* 46: 279–303.
- Kogevinas M. 2001. Human health effects of dioxins: cancer, reproductive and endocrine system effects. *Hum. Reprod. Update* 7: 331–339.
- Kogevinas M, Becher H, Benn T, Bertazzi PA, Boffetta P, Bueno-de-Mesqurta HB, et al. 1997. Cancer mortality in workers exposed to phenoxy herbicides, chlorophenols, and dioxins an expanded and updated international cohort study. *Am. J. Epidemiol.* 145: 1061–1075.
- Krishnan V, Porter W, Santostefano M, Wang X, Safe S. 1995. Molecular mechanism of inhibition of estrogen-induced cathepsin D gene expression by 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in MCF-7 cells. *Mol. Cell. Biol.* 15:6710–6719; doi:10.1128/MCB.15.12.6710.
- Kumle M, Weiderpass E, Braaten T, Persson I, Adami H-O, Lund E. 2002. Use of oral contraceptives and breast cancer risk the Norwegian-Swedish women's lifestyle and Health Cohort Study. *Cancer Epidemiol. Biomarkers Prev.* 11: 1375–1381.
- Lagarde F, Beausoleil C, Belcher SM, Belzunces LP, Emond C, Guerbet M, et al. 2015. Non-monotonic dose-response relationships and endocrine disruptors: a qualitative method of assessment. *Environ. Health Glob. Access Sci. Source* 14:13; doi:10.1186/1476-069X-14-13.
- Lamartiniere CA. 2002. Timing of exposure and mammary cancer risk. *J. Mammary Gland Biol. Neoplasia* 7: 67–76.
- Lambrechtsen J, Gerke O, Egstrup K, Sand NP, Nørgaard BL, Petersen H, et al. 2012. The relation between coronary artery calcification in asymptomatic subjects and both traditional risk factors and living in the city centre: a DanRisk substudy. *J. Intern. Med.* 271:444–450; doi:10.1111/j.1365-2796.2011.02486.x.
- Landi MT, Bergen AW, Baccarelli A, Patterson DG, Grassman J, Ter-Minassian M, et al. 2005. CYP1A1 and CYP1B1 genotypes, haplotypes, and TCDD-induced gene expression in subjects from Seveso, Italy. *Toxicology* 207:191–202; doi:10.1016/j.tox.2004.08.021.
- Latino-Martel P, Cottet V, Druesne-Pecollo N, Pierre FHF, Touillaud M, Touvier M, et al. 2016. Alcoholic beverages, obesity, physical activity and other nutritional factors, and cancer risk: A review of the evidence. *Crit. Rev. Oncol. Hematol.* 99:308–323; doi:10.1016/j.critrevonc.2016.01.002.
- Lauby-Secretan B, Loomis D, Grosse Y, Ghissassi FE, Bouvard V, Benbrahim-Tallaa L, et al. 2013. Carcinogenicity of polychlorinated biphenyls and polybrominated biphenyls. *Lancet Oncol.* 14:287–288; doi:10.1016/S1470-2045(13)70104-9.
- Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K, et al. 2016. Body Fatness and Cancer--Viewpoint of the IARC Working Group. *N. Engl. J. Med.* 375:794–798; doi:10.1056/NEJMSr1606602.
- Lee W-H, Boyer TG. 2001. BRCA1 and BRCA2 in breast cancer. *The Lancet* 358: S5.
- Leijds MM, Koppe JG, Olie K, Aalderen WMC van, Voogt P de, Vulsma T, et al. 2008. Delayed initiation of breast development in girls with higher prenatal dioxin exposure; a longitudinal cohort study. *Chemosphere* 73:999–1004; doi:10.1016/j.chemosphere.2008.05.053.

- Leng L, Li J, Luo X, Kim J, Li Y, Guo X, et al. 2016. Polychlorinated biphenyls and breast cancer: A congener-specific meta-analysis. *Environ. Int.* 88:133–141; doi:10.1016/j.envint.2015.12.022.
- Lew BJ, Collins LL, O'Reilly MA, Lawrence BP. 2009. Activation of the Aryl Hydrocarbon Receptor during Different Critical Windows in Pregnancy Alters Mammary Epithelial Cell Proliferation and Differentiation. *Toxicol. Sci.* 111:151–162; doi:10.1093/toxsci/kfp125.
- Liu R, Nelson DO, Hurley S, Hertz A, Reynolds P. 2015. Residential Exposure to Estrogen Disrupting Hazardous Air Pollutants and Breast Cancer Risk: The California Teachers Study. *Epidemiology* 26:365–373; doi:10.1097/EDE.0000000000000277.
- Loomis D, Grosse Y, Lauby-Secretan B, Ghissassi FE, Bouvard V, Benbrahim-Tallaa L, et al. 2013. The carcinogenicity of outdoor air pollution. *Lancet Oncol.* 14:1262–1263; doi:10.1016/S1470-2045(13)70487-X.
- López-Carrillo L, Hernández-Ramírez RU, Calafat AM, Torres-Sánchez L, Galván-Portillo M, Needham LL, et al. 2009. Exposure to Phthalates and Breast Cancer Risk in Northern Mexico. *Environ. Health Perspect.* 118:539–544; doi:10.1289/ehp.0901091.
- Lucas F, Niravong M, Villemainot S, Kaaks R, Clavel-Chapelon F. 1995. Estimation of food portion size using photographs: validity, strengths, weaknesses and recommendations. *J. Hum. Nutr. Diet.* 8: 65–74.
- M Hall J. 2013. The Aryl-hydrocarbon Receptor (AhR) as a Therapeutic Target in Human Breast Cancer. *J. Steroids Horm. Sci.* 5; doi:10.4172/2157-7536.1000140.
- Mandal PK. 2005. Dioxin: a review of its environmental effects and its aryl hydrocarbon receptor biology. *J. Comp. Physiol. B* 175:221–230; doi:10.1007/s00360-005-0483-3.
- Mantel N. 1973. Synthetic Retrospective Studies and Related Topics. *Biometrics* 29:479–486; doi:10.2307/2529171.
- Manuwald U, Garrido MV, Berger J, Manz A, Baur X. 2012. Mortality study of chemical workers exposed to dioxins: follow-up 23 years after chemical plant closure. *Occup. Environ. Med.* 69: 636–642.
- Marchbanks PA, McDonald JA, Wilson HG, Folger SG, Mandel MG, Daling JR, et al. 2002. Oral contraceptives and the risk of breast cancer. *N. Engl. J. Med.* 346: 2025–2032.
- McCormack VA, Silva I dos S. 2006. Breast Density and Parenchymal Patterns as Markers of Breast Cancer Risk: A Meta-analysis. *Cancer Epidemiol. Biomarkers Prev.* 15:1159–1169; doi:10.1158/1055-9965.EPI-06-0034.
- McPherson K, Steel CM, Dixon JM. 2000. Breast cancer—epidemiology, risk factors, and genetics. *BMJ* 321: 624–628.
- Milbrath MO, Wenger Y, Chang C-W, Emond C, Garabrant D, Gillespie BW, et al. 2009. Apparent Half-Lives of Dioxins, Furans, and Polychlorinated Biphenyls as a Function of Age, Body Fat, Smoking Status, and Breast-Feeding. *Environ. Health Perspect.* 117:417–425; doi:10.1289/ehp.11781.
- Millikan R, DeVoto E, Duell EJ, Tse CK, Savitz DA, Beach J, et al. 2000. Dichlorodiphenyldichloroethene, polychlorinated biphenyls, and breast cancer among African-American and white women in North Carolina. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 9: 1233–1240.
- Molinié F, Vanier A, Woronoff AS, Guizard AV, Delafosse P, Velten M, et al. 2014. Trends in breast cancer incidence and mortality in France 1990–2008. *Breast Cancer Res. Treat.* 147:167–175; doi:10.1007/s10549-014-3073-9.
- Mordukhovich I, Beyea J, Herring AH, Hatch M, Stellman SD, Teitelbaum SL, et al. 2015. Vehicular Traffic-Related Polycyclic Aromatic Hydrocarbon Exposure and Breast Cancer Incidence: The Long Island Breast Cancer Study Project (LIBCSP). *Environ. Health Perspect.* 124; doi:10.1289/ehp.1307736.

- Moysich KB, Ambrosone CB, Vena JE, Shields PG, Mendola P, Kostyniak P, et al. 1998. Environmental organochlorine exposure and postmenopausal breast cancer risk. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 7: 181–188.
- Moysich KB, Shields PG, Freudenheim JL, Schisterman EF, Vena JE, Kostyniak P, et al. 1999. Polychlorinated biphenyls, cytochrome P4501A1 polymorphism, and postmenopausal breast cancer risk. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 8: 41–44.
- Nakao T, Aozasa O, Ohta S, Miyata H. 2005. Survey of human exposure to PCDDs, PCDFs, and coplanar PCBs using hair as an indicator. *Arch. Environ. Contam. Toxicol.* 49:124–130; doi:10.1007/s00244-004-0059-3.
- Nickels S, Truong T, Hein R, Stevens K, Buck K, Horwitz MS, et al. 2013. Evidence of Gene–Environment Interactions between Common Breast Cancer Susceptibility. *Issue PLoS Genet.* 9 3.
- Norman SA, Berlin JA, Soper KA, Middendorf BF, Stolley PD. 1995. Cancer Incidence in a Group of Workers Potentially Exposed to Ethylene Oxide. *Int. J. Epidemiol.* 24:276–284; doi:10.1093/ije/24.2.276.
- Nzihou A, Themelis NJ, Kemiha M, Benhamou Y. 2012. Dioxin emissions from municipal solid waste incinerators (MSWIs) in France. *Waste Manag.* 32:2273–2277; doi:10.1016/j.wasman.2012.06.016.
- Oakes M, Baxter L, Long TC. 2014. Evaluating the application of multipollutant exposure metrics in air pollution health studies. *Environ. Int.* 69:90–99; doi:10.1016/j.envint.2014.03.030.
- Ogawa T, Asai Y, Yamashita M, Takasuga T. 2003. Detectable dioxins in human saliva and their effects on gingival epithelial cells. *J. Dent. Res.* 82: 849–853.
- Ohtake F, Takeyama K, Matsumoto T, Kitagawa H, Yamamoto Y, Nohara K, et al. 2003. Modulation of oestrogen receptor signalling by association with the activated dioxin receptor. *Nature* 423:545–550; doi:10.1038/nature01606.
- Palmer JR, Wise LA, Hatch EE, Troisi R, Titus-Ernstoff L, Strohsnitter W, et al. 2006. Prenatal Diethylstilbestrol Exposure and Risk of Breast Cancer. *Cancer Epidemiol. Biomarkers Prev.* 15:1509–1514; doi:10.1158/1055-9965.EPI-06-0109.
- Park J-H, Cha ES, Ko Y, Hwang M-S, Hong J-H, Lee WJ. 2014. Exposure to Dichlorodiphenyltrichloroethane and the Risk of Breast Cancer: A Systematic Review and Meta-analysis. *Osong Public Health Res. Perspect.* 5:77–84; doi:10.1016/j.phrp.2014.02.001.
- Parzefall W. 2002. Risk assessment of dioxin contamination in human food. *Food Chem. Toxicol.* 40: 1185–1189.
- Peng S, Lü B, Ruan W, Zhu Y, Sheng H, Lai M. 2011. Genetic polymorphisms and breast cancer risk: evidence from meta-analyses, pooled analyses, and genome-wide association studies. *Breast Cancer Res. Treat.* 127:309–324; doi:10.1007/s10549-011-1459-5.
- Pesatori A, Consonni D, Rubagotti M, Grillo P, Bertazzi P. 2009. Cancer incidence in the population exposed to dioxin after the “Seveso accident”: twenty years of follow-up. *Environ. Health* 8:39; doi:10.1186/1476-069X-8-39.
- Petit C, Blangiardo M, Richardson S, Coquet F, Chevrier C, Cordier S. 2012. Association of Environmental Insecticide Exposure and Fetal Growth With a Bayesian Model Including Multiple Exposure Sources: The PELAGIE Mother-Child Cohort. *Am. J. Epidemiol.* 175:1182–1190; doi:10.1093/aje/kwr422.
- PNSE2. 2009. Deuxième plan national Santé-environnement (PNSE2) - 2009 - 2013 - Santé - Ministère des Affaires sociales et de la Santé. Available: <http://social-sante.gouv.fr/ministere/documentation-et-publications-officielles/rapports/sante/article/deuxieme-plan-national-sante-environnement-pnse2-2009-2013> [accessed 3 October 2016].

- Popp JA. 2006. A Weight-of-Evidence Analysis of the Cancer Dose-Response Characteristics of 2,3,7,8-Tetrachlorodibenzodioxin (TCDD). *Toxicol. Sci.* 89:361–369; doi:10.1093/toxsci/kfj016.
- Pronk A, Nuckols JR, De Roos AJ, Airola M, Colt JS, Cerhan JR, et al. 2013. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case–control study. *Environ. Health* 12: 1.
- PRSE2. 2011. DREAL - PRSE2. Available: <http://www.prse2-rhonealpes.fr/> [accessed 3 October 2016].
- Raaschou-Nielsen O, Pavuk M, LeBlanc A, Dumas P, Weber JP, Olsen A, et al. 2005. Adipose organochlorine concentrations and risk of breast cancer among postmenopausal Danish women. *Cancer Epidemiol. Biomarkers Prev.* 14: 67–74.
- Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. 2008. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet Lond. Engl.* 371:569–578; doi:10.1016/S0140-6736(08)60269-X.
- Revich B, Aksel E, Ushakova T, Ivanova I, Zhuchenko N, Klyuev N, et al. 2001. Dioxin exposure and public health in Chapaevsk, Russia. *Chemosphere* 43: 951–966.
- Reynolds P, Hurley SE, Petreas M, Goldberg DE, Smith D, Gilliss D, et al. 2005. Adipose levels of dioxins and risk of breast cancer. *Cancer Causes Control CCC* 16:525–535; doi:10.1007/s10552-004-7840-5.
- Rhomberg LR, Goodman JE. 2012. Low-dose effects and nonmonotonic dose–responses of endocrine disrupting chemicals: Has the case been made? *Regul. Toxicol. Pharmacol.* 64:130–133; doi:10.1016/j.yrtph.2012.06.015.
- Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. 2002. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women’s Health Initiative randomized controlled trial. *JAMA* 288: 321–333.
- Rudel RA, Ackerman JM, Attfield KR, Brody JG. 2014. New Exposure Biomarkers as Tools For Breast Cancer Epidemiology, Biomonitoring, and Prevention: A Systematic Approach Based on Animal Evidence. *Environ. Health Perspect.*; doi:10.1289/ehp.1307455.
- Rudel RA, Attfield KR, Schifano JN, Brody JG. 2007. Chemicals causing mammary gland tumors in animals signal new directions for epidemiology, chemicals testing, and risk assessment for breast cancer prevention. *Cancer* 109:2635–2666; doi:10.1002/cncr.22653.
- Safe S, Lee S-O, Jin U-H. 2013. Role of the Aryl Hydrocarbon Receptor in Carcinogenesis and Potential as a Drug Target. *Toxicol. Sci.* 135:1–16; doi:10.1093/toxsci/kft128.
- Safe SH. 1995. Modulation of gene expression and endocrine response pathways by 2,3,7,8-tetrachlorodibenzo-p-dioxin and related compounds. *Pharmacol. Ther.* 67:247–281; doi:10.1016/0163-7258(95)00017-B.
- Saintot M, Malaveille C, Hautefeuille A, Gerber M. 2004. Interaction between genetic polymorphism of cytochrome P450-1B1 and environmental pollutants in breast cancer risk. *Eur. J. Cancer Prev.* 13: 83–86.
- Schernhammer ES, Kroenke CH, Laden F, Hankinson SE. 2006. Night Work and Risk of Breast Cancer: *Epidemiology* 17:108–111; doi:10.1097/01.ede.0000190539.03500.c1.
- Schug TT, Janesick A, Blumberg B, Heindel JJ. 2011. Endocrine disrupting chemicals and disease susceptibility. *J. Steroid Biochem. Mol. Biol.* 127:204–215; doi:10.1016/j.jsbmb.2011.08.007.
- Seitz HK, Pelucchi C, Bagnardi V, Vecchia CL. 2012. Epidemiology and Pathophysiology of Alcohol and Breast Cancer: Update 2012. *Alcohol Alcohol* 47:204–212; doi:10.1093/alcalc/ags011.
- Silva E, Rajapakse N, Kortenkamp A. 2002. Something from “nothing”--eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. *Environ. Sci. Technol.* 36: 1751–1756.

- Simonds NI, Ghazarian AA, Pimentel CB, Schully SD, Ellison GL, Gillanders EM, et al. 2016. Review of the Gene-Environment Interaction Literature in Cancer: What Do We Know? *Genet. Epidemiol.* 40:356–365; doi:10.1002/gepi.21967.
- Sirot V, Volatier JL, Calamassi-Tran G, Dubuisson C, Menard C, Dufour A, et al. 2009. Core food of the French food supply: second Total Diet Study. *Food Addit. Contam. Part Chem. Anal. Control Expo. Risk Assess.* 26:623–639; doi:10.1080/02652030802695506.
- Sonawane BR. 1995. Chemical contaminants in human milk: an overview. *Environ. Health Perspect.* 103: 197–205.
- Sonnenschein C, Soto AM. 1998. An updated review of environmental estrogen and androgen mimics and antagonists I. *J. Steroid Biochem. Mol. Biol.* 65:143–150; doi:10.1016/S0960-0760(98)00027-2.
- Soto AM, Sonnenschein C. 2010. Environmental causes of cancer: endocrine disruptors as carcinogens. *Nat. Rev. Endocrinol.* 6:363–370; doi:10.1038/nrendo.2010.87.
- Soulhac L, Salizzoni P, Cierco F-X, Perkins R. 2011. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmos. Environ.* 45:7379–7395; doi:10.1016/j.atmosenv.2011.07.008.
- Soulhac L, Salizzoni P, Mejean P, Didier D, Rios I. 2012. The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmos. Environ.* 49:320–337; doi:10.1016/j.atmosenv.2011.11.031.
- Steenland K, Whelan E, Deddens J, Stayner L, Ward E. 2003. Ethylene oxide and breast cancer incidence in a cohort study of 7576 women (United States). *Cancer Causes Control* 14: 531–539.
- Stewart BW, Kleihues P, International Agency for Research on Cancer, eds. 2003. *World cancer report*. IARC Press, Lyon.
- Straif K, Baan R, Grosse Y, Secretan B, El Ghissassi F, Bouvard V, et al. 2007. Carcinogenicity of shift-work, painting, and fire-fighting. *Lancet Oncol.* 8: 1065–1066.
- Svensson BG, Nilsson A, Hansson M, Rappe C, Akesson B, Skerfving S. 1991. Exposure to dioxins and dibenzofurans through the consumption of fish. *N. Engl. J. Med.* 324:8–12; doi:10.1056/NEJM199101033240102.
- Sylvestre M-P, Abrahamowicz M. 2009. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Stat. Med.* 28:3437–3453; doi:10.1002/sim.3701.
- Takekuma M, Saito K, Ogawa M, Matumoto R, Kobayashi S. 2004. Levels of PCDDs, PCDFs and Co-PCBs in human milk in Saitama, Japan, and epidemiological research. *Chemosphere* 54:127–135; doi:10.1016/S0045-6535(03)00670-2.
- Tard A, Gallotti S, Leblanc J-C, Volatier J-L. 2007. Dioxins, furans and dioxin-like PCBs: Occurrence in food and dietary intake in France. *Food Addit. Contam.* 24:1007–1017; doi:10.1080/02652030701317293.
- Taylor VH, Misra M, Mukherjee SD. 2009. Is red meat intake a risk factor for breast cancer among premenopausal women? *Breast Cancer Res. Treat.* 117:1–8; doi:10.1007/s10549-009-0441-y.
- Tehard B, Lahmann PH, Riboli E, Clavel-Chapelon F, the E3N Group. 2004. Anthropometry, breast cancer and menopausal status: Use of repeated measurements over 10 years of follow-up? results of the French E3N women's cohort study. *Int. J. Cancer* 111:264–269; doi:10.1002/ijc.20213.
- Teitelbaum SL, Belpoggi F, Reinlib L. 2015. Advancing research on endocrine disrupting chemicals in breast cancer: Expert panel recommendations. *Reprod. Toxicol.* 54:141–147; doi:10.1016/j.reprotox.2014.12.015.
- Thiebaut ACM, Kipnis V, Chang S-C, Subar AF, Thompson FE, Rosenberg PS, et al. 2007. Dietary Fat and Postmenopausal Invasive Breast Cancer in the National Institutes of Health-AARP Diet and Health Study Cohort. *JNCI J. Natl. Cancer Inst.* 99:451–462; doi:10.1093/jnci/djk094.

- Thompson D, Easton D. 2004. The genetic epidemiology of breast cancer genes. *J. Mammary Gland Biol. Neoplasia* 9:221–236; doi:10.1023/B:JOMG.0000048770.90334.3b.
- Turunen AW, Männistö S, Kiviranta H, Marniemi J, Jula A, Tiittanen P, et al. 2010. Dioxins, polychlorinated biphenyls, methyl mercury and omega-3 polyunsaturated fatty acids as biomarkers of fish consumption. *Eur. J. Clin. Nutr.* 64: 313–323.
- Uemura H, Arisawa K, Hiyoshi M, Satoh H, Sumiyoshi Y, Morinaga K, et al. 2008. PCDDs/PCDFs and dioxin-like PCBs: Recent body burden levels and their determinants among general inhabitants in Japan. *Chemosphere* 73:30–37; doi:10.1016/j.chemosphere.2008.05.066.
- UNEP. 2013. United Nations Environment Programme | Toolkit for Identification and Quantification of Releases of Dioxins, Furans and Other Unintentional POPs. Available: http://toolkit.pops.int/Publish/Main/01_Index.html [accessed 22 August 2016].
- Vacek PM. 1997. Assessing the effect of intensity when exposure varies over time. *Stat. Med.* 16: 505–513.
- Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. 2000. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). *Cancer Causes Control* 11: 653–662.
- Van den Berg M, De Jongh J, Poiger H, Olson JR. 1994. The toxicokinetics and metabolism of polychlorinated dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs) and their relevance for toxicity. *CRC Crit. Rev. Toxicol.* 24: 1–74.
- Van Den Brandt PA, Spiegelman D, Yaun S-S, Adami H-O, Beeson L, Folsom AR, et al. 2000. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am. J. Epidemiol.* 152: 514–527.
- van der Groep P, van der Wall E, van Diest PJ. 2011. Pathology of hereditary breast cancer. *Cell. Oncol. Dordr.* 34:71–88; doi:10.1007/s13402-011-0010-3.
- Van Liere MJ, Lucas F, Clavel F, Slimani N, Villemainot S. 1997. Relative validity and reproducibility of a French dietary history questionnaire. *Int. J. Epidemiol.* 26: S128.
- Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR, Lee D-H, et al. 2012. Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses. *Endocr. Rev.* 33:378–455; doi:10.1210/er.2011-1050.
- Viel J-F, Clément M-C, Hägi M, Grandjean S, Challier B, Danzon A. 2008. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure. *Int. J. Health Geogr.* 7:4; doi:10.1186/1476-072X-7-4.
- Vienneau D, de Hoogh K, Briggs D. 2009. A GIS-based method for modelling air pollution exposures across Europe. *Sci. Total Environ.* 408:255–266; doi:10.1016/j.scitotenv.2009.09.048.
- Viera AJ, Garrett JM, others. 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med* 37: 360–363.
- Wang AT, Vachon CM, Brandt KR, Ghosh K. 2014. Breast Density and Breast Cancer Risk: A Practical Review. *Mayo Clin. Proc.* 89:548–557; doi:10.1016/j.mayocp.2013.12.014.
- Wang F-R, Fang Q-Q, Tang W-M, Xu X-S, Mahapatra T, Mahapatra S, et al. 2015. Nested Case-control Study of Occupational Radiation Exposure and Breast and Esophagus Cancer Risk among Medical Diagnostic X Ray Workers in Jiangsu of China. *Asian Pac. J. Cancer Prev. APJCP* 16: 4699–4704.
- Wang WL, Porter W, Burghardt R, Safe SH. 1997. Mechanism of inhibition of MDA-MB-468 breast cancer cell growth by 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin. *Carcinogenesis* 18: 925–933.
- Warner M, Eskenazi B, Mocarelli P, Gerthoux PM, Samuels S, Needham L, et al. 2002. Serum dioxin concentrations and breast cancer risk in the Seveso Women’s Health Study. *Environ. Health Perspect.* 110: 625.

- Warner M, Mocarelli P, Samuels S, Needham L, Brambilla P, Eskenazi B. 2011. Dioxin exposure and cancer risk in the Seveso Women's Health Study. *Environ. Health Perspect.* 119:1700–1705; doi:10.1289/ehp.1103720.
- White AJ, Bradshaw PT, Herring AH, Teitelbaum SL, Beyea J, Stellman SD, et al. 2016. Exposure to multiple sources of polycyclic aromatic hydrocarbons and breast cancer incidence. *Environ. Int.* 89–90:185–192; doi:10.1016/j.envint.2016.02.009.
- White AJ, Teitelbaum SL, Stellman SD, Beyea J, Steck SE, Mordukhovich I, et al. 2014. Indoor air pollution exposure from use of indoor stoves and fireplaces in association with breast cancer: a case-control study. *Environ. Health* 13: 1.
- White N, teWaterNaude J, van der Walt A, Ravenscroft G, Roberts W, Ehrlich R. 2009. Meteorologically estimated exposure but not distance predicts asthma symptoms in schoolchildren in the environs of a petrochemical refinery: a cross-sectional study. *Environ. Health* 8; doi:10.1186/1476-069X-8-45.
- Wolff MS, Britton JA, Teitelbaum SL, Eng S, Deych E, Ireland K, et al. 2005. Improving Organochlorine Biomarker Models for Cancer Research. *Cancer Epidemiol. Biomarkers Prev.* 14:2224–2236; doi:10.1158/1055-9965.EPI-05-0173.
- Wu Y, Zhang D, Kang S. 2013. Physical activity and risk of breast cancer: a meta-analysis of prospective studies. *Breast Cancer Res. Treat.* 137:869–882; doi:10.1007/s10549-012-2396-7.
- Xu X, Dailey AB, Talbott EO, Ilacqua VA, Kearney G, Asal NR. 2009. Associations of Serum Concentrations of Organochlorine Pesticides with Breast Cancer and Prostate Cancer in U.S. Adults. *Environ. Health Perspect.*; doi:10.1289/ehp.0900919.
- Yoshioka H, Hiromori Y, Aoki A, Kimura T, Fujii-Kuriyama Y, Nagase H, et al. 2012. Possible aryl hydrocarbon receptor-independent pathway of 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced antiproliferative response in human breast cancer cells. *Toxicol. Lett.* 211:257–265; doi:10.1016/j.toxlet.2012.04.005.
- Yu C-L. 2006. Residential Exposure to Petrochemicals and the Risk of Leukemia: Using Geographic Information System Tools to Estimate Individual-Level Residential Exposure. *Am. J. Epidemiol.* 164:200–207; doi:10.1093/aje/kwj182.
- Zhang S, Lei P, Liu X, Li X, Walker K, Kotha L, et al. 2009. The aryl hydrocarbon receptor as a target for estrogen receptor-negative breast cancer chemotherapy. *Endocr. Relat. Cancer* 16:835–844; doi:10.1677/ERC-09-0054.
- Zou B, Wilson JG, Zhan FB, Zeng Y. 2009. Air pollution exposure assessment methods utilized in epidemiological studies. *J. Environ. Monit.* 11:475–490; doi:10.1039/B813889C.

Appendices

Appendix 1. Literature review: list of algorithms

Algorithms used to identify methodological and epidemiological studies that developed a GIS-based metric to estimate exposure to:

- **Dioxins:**

(dioxins[mesh terms] OR dioxin*[tiab] OR tetrachlorodibenzodioxin[mesh terms] OR tetrachlorodibenzo p dioxin[tiab] OR TCDD[tiab] OR polychlorodibenzo-4-dioxin[tiab] OR PCDD[tiab] OR PCDF[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (geographic information systems[mesh terms] OR geographic information system[tiab] OR geograph*[tiab] OR metric[tiab] OR (spatial*[tiab] and (analys*[tiab] OR analyt*[tiab])) OR GIS*[tiab] OR spatial analysis[mesh terms]) AND (plume*[tiab] OR buffer*[tiab] OR distance*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor*[tiab] OR meteorolog*[tiab] OR ((residen*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

- **PAH, dioxin-like compounds and POP:**

(polycyclic hydrocarbons,aromatic[mesh terms] OR polycyclic aromatic hydrocarbon[tiab] OR polychlorinated biphenyls[mesh terms] OR polychlorinated biphenyls[tiab] OR PCB*[tiab] OR PAH*[tiab] OR dioxin-like[tiab] OR persistent organic pollutant*[tiab] OR POP[tiab] OR POPs[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (geographic information systems[mesh terms] OR geographic information system[tiab] OR geograph*[tiab] OR metric[tiab] OR (spatial*[tiab] and (analys*[tiab] OR analyt*[tiab])) OR GIS*[tiab] OR spatial analysis[mesh terms]) AND (plume*[tiab] OR buffer*[tiab] OR distance*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor*[tiab] OR meteorolog*[tiab] OR ((residen*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

- **Air pollution:**

((air pollut*[mesh terms] OR air pollut*[tiab]) OR (environmental pollut*[mesh terms] OR environmental pollut*[tiab]) OR (carbon dioxide[mesh terms] OR carbon dioxide[tiab] OR CO2[tiab]) OR (nitrogen oxides[mesh terms] OR nitrogen oxide*[tiab]) OR (particulate matter[mesh terms] OR particulate matter*[tiab] OR PM[tiab]) OR industrial waste[mesh terms] OR emission*[tiab] OR hazardous waste[mesh terms] OR hazardous waste[tiab] OR industry[mesh terms] OR indust*[tiab] OR incineration[mesh terms] OR incinerat*[tiab] OR combustion[tiab] OR facilit*[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (geographic information systems[mesh terms] OR geographic information system[tiab] OR geograph*[tiab] OR metric[tiab] OR (spatial*[tiab] and (analys*[tiab] OR analyt*[tiab])) OR GIS*[tiab] OR spatial analysis[mesh terms]) AND (plume*[tiab] OR buffer*[tiab] OR distance*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor*[tiab] OR meteorolog*[tiab] OR ((residen*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

Algorithm used to identify studies that modeled dioxin atmospheric dispersion:

(dioxins[mesh terms] OR dioxin*[tiab] OR tetrachlorodibenzodioxin[mesh terms] OR tetrachlorodibenzo p dioxin[tiab] OR TCDD[tiab] OR polychlorodibenzo-4-dioxin[tiab] OR PCDD[tiab] OR PCDF[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (((dispersion[tiab] AND model*[tiab]) OR dispersion model*[tiab]) OR (normal distribution[mesh terms] OR gaussian[tiab] OR gaussian model*[tiab]) OR (ADMS[tiab] OR atmospheric dispersion modeling system[tiab]) OR (lagrangian[tiab] OR lagrangian model*[tiab]) OR air movements[mesh terms]) AND (plume*[tiab] OR buffer*[tiab] OR distance*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor*[tiab] OR meteorolog*[tiab] OR ((residen*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

Appendix 2. Literature review: presentation of 7 articles selected

Table 3. Presentation of the 7 articles selected through the literature review carried out to define the GIS-based metric.

Author and year of publication	Title	Objectives	Method: parameters
Pronk, 2013	Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma (NHL): a case-control study	To evaluate residence near industrial combustion facilities and estimates of dioxin emissions in relation to NHL risk in the United States	Distance decline of $1/d^2$; residential history (15-year period); no meteorological data; no dietary data
Gulliver and Briggs, 2011	STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment	To describe the methods used to develop the STEMS-Air model in a GIS and related data sets To present the results from a validation study using routinely available monitoring data To make observations on potential improvements to the approach for further development work	Gaussian dispersion modeling simplified and considering wind (divided into 45° angles); moving window approach; buffer size around subjects depending on the prevailing wind (FOCALSUM tool in ArcGIS); hourly meteorological parameters (wind speed, wind direction, cloud cover); height of the release of the emissions
Vienneau, 2009	A GIS-based method for modeling air pollution exposures across Europe	To describe, demonstrate and test a moving window technique to derive a distance weighted model of relationships between area-level emissions and monitored concentrations at a set of training sites and apply it to estimate concentrations at unsampled locations across EU	Moving window approach (FOCALSUM tool in ArcGIS); 45° wind direction angles; distance decline of $1/d^2$ and adjusted for measurements; maximum distance of 11 km as a limitation of detectable contributions; wind speed (mean annual; <1 m/s; $1-2.5$ m/s; >2.5 m/s)
White, 2009	Meteorologically estimated exposure but not distance predicts asthma symptoms in schoolchildren in the environs of a petrochemical refinery: a cross-sectional study	To determine whether asthma symptom prevalences were elevated compared to comparable areas in Cape Town and whether there was an association between asthma symptom prevalences and the derived exposure metric	Wind direction and speed (data for the year prior to the study and from the meteorological station closest to the source); manual geocoding of home addresses; GIS to compute the distance from the home address to the midpoint of the source
Zou, 2009	Air pollution exposure assessment methods utilized in epidemiological studies (review)	To outline the methods and criteria utilized in understanding the current state of air pollution exposure assessment; To provide an overview of concepts and research frameworks utilized in air pollution exposure assessment; To review exposure models and biomarkers (developments, applications, merits and weaknesses)	Proximity models: residential distance to receptor; distance-weighted traffic density; geographic information system including buffer analysis, geocoding, track analysis; land use regression including traffic density, industrial land use, open space, location; meteorological and topographical conditions; Air dispersion models: incorporate emission data, basic meteorological data (wind speed and direction),

Author and year of publication	Title	Objectives	Method: parameters
Yu, 2006	Residential exposure to petrochemicals and the risk of leukemia: using geographic information system tools to estimate individual-level residential exposure	To offer perspectives about future development of air pollution exposure assessment methods	<p>topography, ambient temperature, atmospheric turbulence, ambient background pollution concentrations;</p> <p>Hybrid models: combination of personal monitoring with regional monitoring; air dispersion modelling with regional monitoring or personal monitoring, regional grid with local plume models;</p> <p>Human inhalation models: human activity, physiology, chemical and environmental conditions;</p> <p>Biomarkers: biomarkers of susceptibility, of effect, of exposure</p>
Hoek, 2001	Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality	<p>To investigate the relationship between traffic-related air pollution and mortality within the Netherlands Cohort Study</p> <p>To develop methods for exposure assessment and evaluate the contrast in exposure to air pollution in the cohort within a random sub-cohort of 5,000 subjects</p>	<p>Residential history; geocoding of the residences and locations of petrochemical plants; 3 km radius from the geographic centroid of the plants; monthly prevailing wind direction; 90-degrees areas (+/- 45-degree angle from the downwind direction); distance decline of 1/d</p> <p>Residential addresses in a GIS (not the exact address)</p> <p>Exposure=regional background+ urban background+ local background</p> <p>Regional background: distance decline from home to the site= 1/d; 1/d²; 1/d³ + interpolated concentration from network sites to the residential addresses;</p> <p>Urban background: address density (measure the number of addresses) as a proxy for the intensity of human activities in an area;</p> <p>Local background: distance to different road types (<50 m from a major urban road= exposed; >50 m = not exposed)</p>

Appendice 3. GIS-based metric tests: weighted kappa coefficients and 95% CI

Table 4. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested, in Lyon, in 1996, 2002 and 2008.

Table 5. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested, in Le Bugey, in 1996, 2002 and 2008.

Table 6. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested including stack height (h) and exhaust smoke velocity (v), in Lyon, in 1996, 2002 and 2008.

Table 7. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested including stack height (h) and exhaust smoke velocity (v), in Le Bugey, in 1996, 2002 and 2008.

Table 4. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested, in Lyon, in 1996, 2002 and 2008.

Distance decline	Prevailing wind frequency – segment sizes											
	No prevailing wind frequency		90°		30°		10°		10° (4-times weighted) and adjacent segments		10° (2-times weighted) and adjacent segments	
	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	wk (95% CI)	
1/d												
1996	0.64 (0.59-0.69)	0.65 (0.59-0.70)	0.57 (0.51-0.63)	0.64 (0.59-0.69)	0.56 (0.50-0.63)	0.56 (0.50-0.62)						
2002	0.61 (0.54-0.68)	0.63 (0.55-0.70)	0.61 (0.54-0.68)	0.60 (0.52-0.67)	0.60 (0.52-0.67)	0.60 (0.53-0.68)						
2008	0.71 (0.61-0.82)	0.48 (0.34-0.63)	0.56 (0.42-0.70)	0.69 (0.58-0.80)	0.67 (0.56-0.78)	0.69 (0.59-0.80)						
1/d²												
1996	0.61 (0.55-0.66)	0.70 (0.66-0.75)	0.76 (0.72-0.81)	0.75 (0.71-0.79)	0.71 (0.67-0.76)	0.71 (0.67-0.76)						
2002	0.60 (0.53-0.68)	0.71 (0.65-0.78)	0.80 (0.75-0.85)	0.83 (0.76-0.88)	0.82 (0.78-0.87)	0.84 (0.79-0.88)						
2008	0.64 (0.52-0.75)	0.62 (0.49-0.74)	0.66 (0.54-0.77)	0.83 (0.75-0.90)	0.79 (0.70-0.88)	0.81 (0.72-0.89)						
e^{-d}												
1996	0.52 (0.46-0.59)	0.63 (0.56-0.67)	0.63 (0.57-0.68)	0.39 (0.32-0.46)	0.65 (0.58-0.70)	0.65 (0.60-0.70)						
2002	0.52 (0.43-0.60)	0.65 (0.57-0.72)	0.68 (0.60-0.75)	0.32 (0.23-0.42)	0.65 (0.58-0.73)	0.66 (0.59-0.74)						
2008	0.58 (0.45-0.71)	0.60 (0.47-0.73)	0.62 (0.49-0.74)	0.43 (0.27-0.58)	0.71 (0.61-0.81)	0.71 (0.61-0.81)						
e^{-d}/d												
1996	0.48 (0.41-0.55)	0.57 (0.51-0.63)	0.59 (0.53-0.65)	0.64 (0.59-0.69)	0.59 (0.54-0.65)	0.59 (0.54-0.65)						
2002	0.52 (0.42-0.61)	0.63 (0.54-0.71)	0.65 (0.57-0.73)	0.60 (0.52-0.67)	0.65 (0.57-0.73)	0.64 (0.56-0.72)						
2008	0.58 (0.46-0.70)	0.54 (0.40-0.68)	0.60 (0.47-0.73)	0.69 (0.58-0.80)	0.62 (0.49-0.74)	0.62 (0.50-0.74)						
1/d^{1.5}												
1996	0.63 (0.57-0.68)	0.71 (0.66-0.75)	0.70 (0.65-0.75)	0.72 (0.68-0.76)	0.66 (0.61-0.70)	0.66 (0.61-0.71)						
2002	0.61 (0.54-0.68)	0.73 (0.67-0.79)	0.74 (0.68-0.80)	0.75 (0.69-0.81)	0.73 (0.67-0.79)	0.73 (0.67-0.79)						
2008	0.66 (0.54-0.77)	0.58 (0.46-0.70)	0.62 (0.49-0.75)	0.85 (0.78-0.92)	0.79 (0.70-0.88)	0.79 (0.70-0.88)						
1/d³												
1996	0.52 (0.46-0.58)	0.64 (0.58-0.69)	0.67 (0.62-0.72)	0.69 (0.65-0.74)	0.67 (0.62-0.72)	0.68 (0.63-0.73)						
2002	0.55 (0.47-0.64)	0.67 (0.60-0.74)	0.77 (0.70-0.83)	0.77 (0.70-0.83)	0.75 (0.69-0.81)	0.74 (0.68-0.81)						
2008	0.58 (0.45-0.71)	0.60 (0.46-0.74)	0.62 (0.48-0.75)	0.66 (0.54-0.77)	0.69 (0.58-0.81)	0.69 (0.58-0.81)						
e^{-d}*d												
1996	0.55 (0.49-0.61)	0.64 (0.59-0.69)	0.67 (0.63-0.72)	0.24 (0.16-0.32)	0.69 (0.65-0.74)	0.69 (0.64-0.73)						
2002	0.55 (0.47-0.63)	0.68 (0.62-0.75)	0.74 (0.68-0.81)	0.16 (0.06-0.27)	0.74 (0.68-0.80)	0.74 (0.68-0.80)						
2008	0.60 (0.47-0.72)	0.62 (0.50-0.74)	0.69 (0.58-0.81)	0.31 (0.15-0.47)	0.69 (0.60-0.79)	0.69 (0.59-0.79)						
Including wind speed (w)												
1996												0.68 (0.63-0.73)
2002												0.59 (0.52-0.66)
2008												0.71 (0.61-0.81)

Table 5. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested, in Le Bugey, in 1996, 2002 and 2008.

Distance decline	Prevailing wind frequency – segment sizes					
	No prevailing wind frequency wk (95% CI)	90° wk (95% CI)	30° wk (95% CI)	10° wk (95% CI)	10° (4-times weighted) and adjacent segments wk (95% CI)	10° (2-times weighted) and adjacent segments wk (95% CI)
1/d						
1996	0.64 (0.56-0.72)	0.58 (0.50-0.67)	0.63 (0.55-0.70)	0.65 (0.58-0.72)	0.66 (0.59-0.73)	0.63 (0.55-0.70)
2002	0.78 (0.73-0.84)	0.58 (0.51-0.66)	0.59 (0.51-0.70)	0.58 (0.50-0.67)	0.58 (0.49-0.66)	0.58 (0.50-0.65)
2008	0.68 (0.60-0.75)	0.51 (0.42-0.60)	0.53 (0.44-0.63)	0.53 (0.45-0.62)	0.53 (0.44-0.61)	0.52 (0.43-0.61)
1/d²						
1996	0.59 (0.51-0.68)	0.67 (0.59-0.74)	0.73 (0.67-0.80)	0.75 (0.69-0.81)	0.76 (0.71-0.84)	0.79 (0.73-0.85)
2002	0.74 (0.68-0.80)	0.78 (0.72-0.83)	0.84 (0.79-0.89)	0.83 (0.77-0.88)	0.83 (0.78-0.88)	0.82 (0.76-0.87)
2008	0.64 (0.57-0.72)	0.67 (0.60-0.74)	0.72 (0.65-0.78)	0.73 (0.67-0.80)	0.74 (0.68-0.80)	0.73 (0.66-0.79)
e^d						
1996	0.50 (0.40-0.60)	0.56 (0.47-0.65)	0.63 (0.54-0.71)	0.64 (0.56-0.72)	0.65 (0.57-0.73)	0.65 (0.57-0.73)
2002	0.65 (0.58-0.72)	0.73 (0.67-0.80)	0.77 (0.71-0.83)	0.78 (0.72-0.84)	0.78 (0.72-0.84)	0.78 (0.72-0.84)
2008	0.58 (0.50-0.66)	0.67 (0.60-0.74)	0.66 (0.59-0.73)	0.66 (0.59-0.73)	0.67 (0.60-0.74)	0.67 (0.60-0.74)
e^d/d						
1996	0.45 (0.35-0.55)	0.49 (0.39-0.59)	0.51 (0.42-0.60)	0.54 (0.45-0.63)	0.53 (0.44-0.61)	0.53 (0.44-0.61)
2002	0.63 (0.56-0.71)	0.66 (0.58-0.73)	0.73 (0.66-0.79)	0.74 (0.67-0.81)	0.73 (0.66-0.80)	0.73 (0.67-0.80)
2008	0.57 (0.48-0.65)	0.60 (0.52-0.68)	0.59 (0.51-0.67)	0.62 (0.54-0.69)	0.62 (0.54-0.69)	0.62 (0.54-0.69)
1/d^{1.5}						
1996	0.61 (0.53-0.69)	0.66 (0.58-0.74)	0.73 (0.67-0.80)	0.72 (0.65-0.78)	0.73 (0.67-0.80)	0.73 (0.67-0.80)
2002	0.78 (0.72-0.83)	0.69 (0.63-0.75)	0.78 (0.73-0.84)	0.74 (0.68-0.80)	0.76 (0.70-0.82)	0.74 (0.68-0.80)
2008	0.68 (0.61-0.76)	0.62 (0.54-0.69)	0.64 (0.56-0.72)	0.66 (0.58-0.73)	0.66 (0.58-0.73)	0.65 (0.58-0.72)
1/d³						
1996	0.51 (0.42-0.60)	0.58 (0.49-0.66)	0.69 (0.62-0.77)	0.70 (0.63-0.77)	0.72 (0.64-0.79)	0.72 (0.65-0.79)
2002	0.68 (0.61-0.76)	0.78 (0.71-0.84)	0.83 (0.77-0.88)	0.83 (0.78-0.89)	0.85 (0.80-0.90)	0.85 (0.80-0.90)
2008	0.62 (0.54-0.69)	0.68 (0.61-0.75)	0.68 (0.61-0.74)	0.67 (0.60-0.73)	0.69 (0.62-0.76)	0.69 (0.62-0.76)
e^d*d						
1996	0.53 (0.44-0.63)	0.66 (0.60-0.75)	0.70 (0.63-0.77)	0.75 (0.68-0.82)	0.75 (0.68-0.82)	0.75 (0.68-0.82)
2002	0.65 (0.58-0.72)	0.74 (0.68-0.80)	0.78 (0.73-0.84)	0.80 (0.75-0.85)	0.82 (0.77-0.87)	0.81 (0.76-0.86)
2008	0.60 (0.52-0.68)	0.64 (0.57-0.71)	0.68 (0.62-0.75)	0.68 (0.62-0.75)	0.68 (0.62-0.75)	0.68 (0.62-0.75)
Including wind speed (w)						
1996						0.76 (0.70-0.82)
2002						0.58 (0.49-0.66)
2008						0.59 (0.51-0.67)

Table 6. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested including stack height (h) and exhaust smoke velocity (v), in Lyon, in 1996, 2002 and 2008.

Stack height (h) and exhaust smoke velocity (v) combinations		Prevailing wind frequency – segment sizes	
		10°	10° (2-times weighted) and adjacent segments
		wk (95% CI)	wk (95% CI)
1/h	1996	0.68 (0.63-0.72)	0.64 (0.59-0.69)
	2002	0.77 (0.70-0.84)	0.77 (0.70-0.83)
	2008	0.75 (0.65-0.85)	0.75 (0.64-0.86)
1/√h	1996	0.72 (0.68-0.76)	0.68 (0.64-0.73)
	2002	0.80 (0.75-0.87)	0.81 (0.77-0.86)
	2008	0.83 (0.75-0.91)	0.79 (0.70-0.88)
1/h²	1996	0.60 (0.54-0.65)	0.57 (0.51-0.64)
	2002	0.69 (0.61-0.77)	0.70 (0.62-0.78)
	2008	0.64 (0.49-0.78)	0.56 (0.42-0.70)
1/(h*v)	1996	0.69 (0.65-0.74)	0.67 (0.62-0.71)
	2002	0.78 (0.72-0.84)	0.78 (0.72-0.84)
	2008	0.77 (0.68-0.86)	0.81 (0.72-0.90)
1/(√h*v)	1996	0.73 (0.69-0.77)	0.70 (0.66-0.75)
	2002	0.82 (0.76-0.87)	0.82 (0.77-0.87)
	2008	0.81 (0.72-0.89)	0.81 (0.73-0.89)
1/(h²*v)	1996	0.62 (0.57-0.68)	0.60 (0.54-0.66)
	2002	0.71 (0.63-0.79)	0.71 (0.63-0.78)
	2008	0.66 (0.53-0.78)	0.66 (0.53-0.78)

Table 7. Weighted kappa (wk) coefficients and 95% CI for agreement between dioxin concentration modeling and estimations of the GIS-based metrics tested including stack height (h) and exhaust smoke velocity (v), in Le Bugey, in 1996, 2002 and 2008.

Stack height (h) and exhaust smoke velocity (v) combinations		Prevailing wind frequency – segment sizes	
		10° wk (95% CI)	10° (2-times weighted) and adjacent segments wk (95% CI)
1/h	1996	0.68 (0.60-0.75)	0.70 (0.63-0.77)
	2002	0.77 (0.71-0.83)	0.78 (0.72-0.84)
	2008	0.73 (0.67-0.80)	0.73 (0.66-0.79)
1/√h	1996	0.73 (0.66-0.79)	0.73 (0.66-0.79)
	2002	0.79 (0.73-0.85)	0.80 (0.74-0.86)
	2008	0.75 (0.69-0.81)	0.76 (0.70-0.82)
1/h²	1996	0.54 (0.45-0.63)	0.55 (0.46-0.64)
	2002	0.64 (0.57-0.72)	0.64 (0.56-0.72)
	2008	0.63 (0.55-0.70)	0.61 (0.53-0.69)
1/(h*v)	1996	0.70 (0.62-0.78)	0.71 (0.64-0.78)
	2002	0.78 (0.72-0.84)	0.81 (0.75-0.86)
	2008	0.74 (0.68-0.80)	0.73 (0.66-0.79)
1/(√h*v)	1996	0.74 (0.68-0.81)	0.77 (0.71-0.83)
	2002	0.79 (0.73-0.85)	0.81 (0.75-0.86)
	2008	0.73 (0.67-0.80)	0.73 (0.67-0.80)
1/(h²*v)	1996	0.59 (0.51-0.68)	0.57 (0.48-0.65)
	2002	0.67 (0.59-0.74)	0.70 (0.63-0.77)
	2008	0.65 (0.57-0.73)	0.63 (0.56-0.71)

Appendice 4. Questionnaire on lifetime residential history



Questionnaire Histoire résidentielle



Villejuif, octobre 2013

Chère Madame,

Depuis de nombreuses années, vous participez à l'étude E3N et nous vous en sommes très reconnaissants.

Nous vous sollicitons aujourd'hui pour le remplissage d'un questionnaire sur les effets des dioxines sur la santé. Les dioxines sont des perturbateurs hormonaux produits lors de processus de combustion (industriels, domestiques ou accidentels) et ils s'accumulent dans la chaîne alimentaire. Il est important pour nous de recueillir des informations sur les caractéristiques de vos logements successifs, leur environnement immédiat ainsi que sur certaines de vos habitudes alimentaires.

Nous avons donc besoin de votre aide pour retracer, de la manière la plus fine possible, votre histoire résidentielle depuis votre naissance. Pour cette étude, il est important que vous puissiez compléter toutes les adresses où vous avez séjourné pendant au moins un an. Nous nous intéressons également à la localisation des lieux de scolarisation et de travail sur lesquels vous avez passé au moins une année.

Afin d'éviter les ratures et de respecter l'ordre chronologique, nous vous invitons à vous remémorer votre histoire résidentielle sur papier libre avant de remplir ce questionnaire. Nous vous remercions de bien vouloir **écrire en majuscules d'imprimerie avec un stylo noir** et de ne pas déborder des cases. Si vous le jugez nécessaire, vous avez la possibilité de cocher plusieurs cases par question.

Veillez par avance accepter nos plus sincères remerciements pour votre précieuse collaboration.

Françoise Clavel-Chapelon
Directrice de Recherche
Responsable de l'enquête E3N

Docteur Béatrice Fervers
Professeur associé
Responsable Unité Cancer et Environnement
Centre Léon Bérard, Lyon

1. Historique de vos lieux de résidence depuis votre naissance jusqu'à aujourd'hui

N'indiquez que les lieux de résidence où vous avez passé au moins une année entière.

Commencez par le lieu où votre mère habitait pendant sa grossesse, puis continuez par les différentes adresses où vous avez habité, par ordre chronologique : 1,2,3, etc.

	Année		Adresse		
	Début	Fin	Numéro	Nom de la voie	Complément (dont étage)
0. Pendant la grossesse de votre mère					
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					

1. Historique de vos lieux de résidence depuis votre naissance jusqu'à aujourd'hui

Adresses	En quelle année ce logement a-t-il été construit ?				Ce logement donnait-il sur cour ou sur rue ? Si logement sur rue, indiquez l'intensité du trafic					Quel mode de chauffage utilisez-vous ?				
	A	B	C	X	A	B	C	D	X	A	B	C	D	X
0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

1. Historique de vos lieux de résidence depuis votre naissance jusqu'à aujourd'hui

Adresses	A quelle fréquence consommez-vous des produits cuisinés au barbecue ?					Consommez-vous des denrées alimentaires											
	A = jamais B = rarement (1 à 2 fois par an) C = occasionnellement D = fréquemment (au moins 1 fois par mois) X = ne sais pas					- les produits de votre production - les produits issus de la cueillette - les préparations réalisées à partir de produits de votre production - les achats ou dons											
						Fruits et/ou légumes						Oeufs					
	A	B	C	D	X	A	B	C	D	E	X	A	B	C	D	E	X
0	<i>Sans objet</i>																
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Historique de vos lieux de scolarisation, de formation et de travail (depuis la maternelle)

N'indiquez que les lieux où vous avez passé au moins une année entière.

	Année		Sans activité professionnelle *	Nom de l'établissement ou du lieu de travail	Numéro	Nom de la voie
	Début	Fin				
I			<input type="checkbox"/>			
II			<input type="checkbox"/>			
III			<input type="checkbox"/>			
IV			<input type="checkbox"/>			
V			<input type="checkbox"/>			
VI			<input type="checkbox"/>			
VII			<input type="checkbox"/>			
VIII			<input type="checkbox"/>			
IX			<input type="checkbox"/>			
X			<input type="checkbox"/>			
XI			<input type="checkbox"/>			
XII			<input type="checkbox"/>			
XIII			<input type="checkbox"/>			
XIV			<input type="checkbox"/>			
XV			<input type="checkbox"/>			
XVI			<input type="checkbox"/>			
XVII			<input type="checkbox"/>			

* Si, pendant une période donnée, vous étiez femme au foyer, cochez la case et complétez la suite de votre histoire professionnelle sur la ligne en-dessous.

2. Historique de vos lieux de scolarisation, de formation et de travail (depuis la maternelle)

Adresses	Combien de temps par jour passiez-vous sur ce lieu ?					Combien de temps mettiez-vous pour vous y rendre ?			
	A	B	C	D	X	A	B	C	X
I	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
II	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
III	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
IV	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
VI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
VII	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
VIII	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
IX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
XI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
XII	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
XIII	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
XIV	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
XV	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
XVI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
XVII	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendice 5. Results of the analyses of the QRH in the Rhône-Alpes region

Study population

Among the 1,642 women who received the QRH, 1,027 completed it by May 2015 for a return rate of 68%. There were 30 deaths and abandonments, 61 refusals (the main reasons were: questionnaire too complicated, no time, memory problems and eventful residential history), 5 return-to-senders due to wrong addresses and 519 non-respondents. After an initial cleaning that consisted in identifying non-filled and unusable questionnaires, analyses of outliers and missing data were performed on 973 questionnaires for the residential history and 977 questionnaires for the occupational history. The QRH's response rate was 57% (Figure 11).

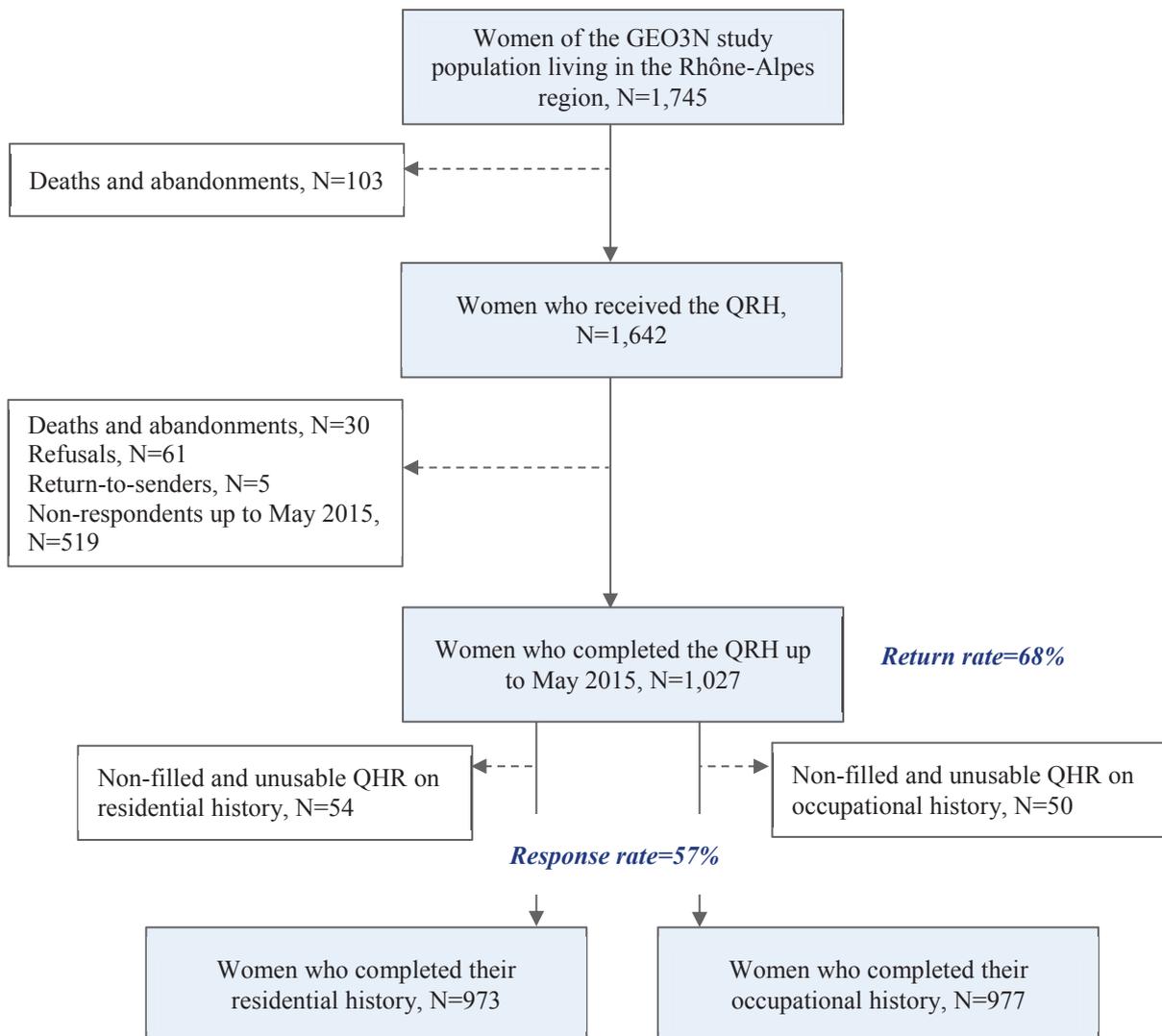


Figure 11. Flowchart of the selection of QRH for quality analysis in the Rhône-Alpes region, France. GEO3N project.

Quality and contents of the filling

A total of 8,549 residential addresses and 7,576 occupational addresses were recorded from the QRH analyzed. On average on lifetime, 8 places of residence and 8 workplaces have been recorded for each woman. Few errors in the filling of the QRH have been identified (Table 8). Less than 1% of

addresses were not chronological; this was explained by the fact that woman could have lived in more than one home during the same period, generally after retirement. For 16.3% of homes, the activity of burning green waste was reported without there being any garden, which meant that burning green waste could have happened outside the garden. For the occupational history, a box could be checked when the woman did not have any work activity (pregnancy, unemployment, retirement); in that case, no address was to be completed and it was assigned the residential address of that period. However, some respondents also checked that box when they were student and provided an address (7% of occupational addresses); that resulted in the non-utilization of that data in the analysis.

The QRH had an overall filling rate of 93%; most of the women completed the questionnaire fully. The percentage of missing data was higher for the first residential and occupational addresses, which are the older ones and require a greater effort of memory. Ten percent of starting and ending years were missing for residential addresses and 7.6% for occupational addresses (Table 8). Between 7.6% and 9.8% of data on consumption frequency of locally produced foods were missing, probably because of the difficulty remembering eating habits at each residence and for the lifetime. Moreover, the modality “Do not know” was proposed but rarely checked.

Table 8. Frequency (n) and percentage (%) of missing data and outliers of the QRH analyzed in May 2015. GEO3N project.

	Missing data	Outliers
	n (%)	n (%)
Residential addresses, N=8549:		
Starting and ending years for residential addresses	861 (10.0)	29 (0.3)
Construction period	226 (2.6)	79 (0.9)
View from the home	281 (3.3)	-
Heating system	281 (3.3)	-
Use of a stove or fireplace	1007 (11.8)	-
Garden	498 (5.8)	49 (0.6)
Burning of green waste	500 (5.8)	1397 (16.3)
Quantity of green waste burnt	797 (9.3)	701 (8.1)
Barbecue cooking	247 (3.3)	33 (0.4)
Fruit and vegetable consumption	576 (7.6)	98 (1.3)
Egg consumption	702 (9.3)	45 (0.6)
Dairy product consumption	710 (9.4)	65 (0.9)
Meat consumption	746 (9.8)	40 (0.5)
Fish and seafood consumption	722 (9.5)	37 (0.5)
Occupational addresses, N=7576:		
Starting and ending years for occupational addresses	632 (7.6)	12 (0.1)
Home-to-work journey time	569 (6.8)	73 (0.9)
Time spent at work	315 (3.8)	107 (0.1)
Transportation	592 (7.1)	-

More than half of the homes have been built before 1948 (Table 9). Heating with wood or coal represented 32.6% of the domestic heating. In 54.8% of the homes, no stove, wood stove or fireplace was used. Women reported to have a garden in 45.2% of their homes, and for 69.9% of them, there was no burning of green waste; quantity of green waste burnt was not known in for 68.7%. In 71.8% of residencies, there was no barbecue cooking (Table 10). Women consumed very often or frequently fruits and vegetables produced locally in 43.1% of the homes. In contrast, eggs, dairy products, meat and seafood were never consumed by women in 34.8%, 42.2%, 44.4% and 55.9% of the homes, respectively. However, these observations must be tempered due to the sizeable percentage of missing data.

Table 9. Descriptive analysis of the lifetime residential addresses from the QRH (excluding missing data and outliers). GEO3N Project. (1/2).

	n (%)
Construction period, N=8244	
Before 1948	4252 (51.6)
Between 1948 and 1974	2547 (30.8)
After 1974	938 (11.4)
Do not know	507 (6.1)
View from the home, N=8549	
Courtyard and garden	4346 (50.8)
Low-traffic street	2165 (25.3)
Moderate-traffic street	1497 (17.5)
Intensive-traffic street	541 (6.3)
Heating system, N=7982	
Individual heating with wood or coal	2606 (32.6)
Other individual heating system	2433 (30.4)
Collective central heating	2296 (28.8)
No heating system	226 (2.8)
Do not know	421 (5.3)
Use of a stove or fireplace, N=7398	
Stove or wood stove	2280 (30.8)
Open fireplace	356 (4.8)
Closed fireplace	253 (3.4)
No use	4054 (54.8)
Do not know	455 (6.2)
Garden, N=8002	
Yes	3615 (45.2)
No	4387 (54.8)
Burning of green waste, N=5038	
Never	2528 (50.2)
Rarely	896 (17.8)
Occasionally	721 (14.3)
Frequently	88 (1.7)
Do not know	805 (16.0)
Quantity of green waste burnt, N=2463	
Less than 1 m ³	686 (27.9)
From 1 m ³ to 5 m ³	521 (21.2)
More than 5 m ³	84 (3.4)
Do not know	1172 (47.6)

Table 10. Descriptive analysis of the lifetime residential addresses from the QRH (excluding missing data and outliers). GEO3N Project. (2/2).

	n (%)
Barbecue cooking frequency, N=7296	
Never	5237 (71.8)
Rarely	937 (12.8)
Occasionally	820 (12.3)
Frequently	145 (2.0)
Do not know	157 (2.2)
Fruit and vegetable consumption frequency, N=6902	
Never	1790 (25.9)
Rarely	647 (9.4)
Occasionally	932 (13.5)
Frequently	870 (12.6)
Very often	2103 (30.5)
Do not know	560 (8.1)
Egg consumption frequency, N=6802	
Never	2368 (34.8)
Rarely	541 (6.3)
Occasionally	917 (10.7)
Frequently	783 (9.1)
Very often	1582 (18.4)
Do not know	638 (7.4)
Dairy product consumption frequency, N=6801	
Never	2869 (42.2)
Rarely	488 (7.2)
Occasionally	706 (10.4)
Frequently	539 (6.3)
Very often	1566 (18.2)
Do not know	633 (7.4)
Meat consumption frequency, N=6790	
Never	3016 (44.4)
Rarely	642 (9.5)
Occasionally	748 (11.0)
Frequently	574 (8.5)
Very often	1049 (15.4)
Do not know	761 (11.2)
Fish and seafood consumption frequency, N=6817	
Never	3808 (55.9)
Rarely	705 (10.3)
Occasionally	716 (10.5)
Frequently	414 (6.1)
Very often	514 (7.5)
Do not know	660 (9.7)

For 65.5% of the education and work places, women took less than 30 minutes to get there, on foot or by bike for 56.0% of them (Table 11), and most of them (72.8%) stayed for up to 8 hours.

Table 11. Descriptive analysis of the lifetime occupational addresses from the QRH (excluding missing data and outliers). GEO3N Project.

	n (%)
Home-to-work journey time, N=7717	
Less than 30 minutes	5058 (65.5)
From 30 min to 1 hour	2050 (26.6)
More than 1 hour	518 (6.7)
Do not know	91 (1.2)
Time spent at work by day, N=7937	
Less than 6 hours	2744 (34.6)
From 6 to 8 hours	3032 (38.2)
From 8 to 10 hours	978 (12.3)
More than 10 hours	1105 (13.9)
Do not know	78 (1.0)
Transportation, N=7767	
On foot or by bike	4351 (56.0)
Motorized wheels	152 (2.0)
Diesel vehicle, gasoline vehicle or public transportation	3227 (41.5)
Do not know	37 (0.5)

Geocoding of the lifetime residential and occupational addresses

From the QRH completed, the addresses have been entered by the partner company Vivetic that provided two databases containing 8,855 records for the residential addresses and 8,536 records for occupational addresses. Addresses were prepared for geocoding, which allowed identification of addresses that could not be geocoded: addresses outside France; addresses without zip code and the name of the town; addresses without zip code and whose name of the town exists in several counties; addresses without the name of the town and whose zip code exists for several towns. Names of schools were also harmonized to facilitate automatic geocoding.

Geocoding of the lifetime residential and occupational addresses of the women living in the Rhône-Alpes region was performed by a geocoding technician (CCa) using the ArcGIS 10.0 Software (ESRI, Redlands, CA, USA) (Article #2). Because the women from the E3N cohort were mostly teachers, the georeferenced database listing the first and second degrees' French schools in 2014, provided by the Education Nationale, was used to geocode occupational addresses (Education Nationale). An accuracy level was assigned to each address geocoded: 1, exact address; 2, interpolated address; 3, street segment; 4, locality; 5, town hall; 6, zip code; 7: city. To improve accuracy, a manual geocoding completed the automatic geocoding for addresses with an accuracy level at 2 and more and for addresses that were not localized to the right zip code.

A total of 8,226 (92.9%) residential addresses were geocoded; 629 (7.1%) addresses could not be geocoded because they were incomplete or outside of France; 5,597 (63.2%) addresses were geocoded automatically and 2,629 (29.7%) manually. Residential addresses were localized at the address (51.0%), at the street segment (27.7%), at the locality (9.6%), at the county-seat (11.5%) and at the city (0.2%). Among the 8,536 occupational addresses, 7,920 (92.8%) were geocoded; 616 (7.2%) could not be geocoded; 6,328 (74.1%) were geocoded automatically and 1,592 (18.7%) manually. Accuracy of geocoding of the occupational addresses was 54.8% at the address, 23.9% at the street segment, 1.6% at the locality, 19.0% at the county-seat and 0.7% at the city.

Rural or urban status of the lifetime residential and occupational addresses

A “rural” or “urban” status was assigned (by CCA) to each address of the residential history and workplaces that were geocoded. According to the National Geographic Information Institute (IGN) and the National Research Institute of Demography and Population (INED), an urban commune was defined as a commune with more than 2,000 inhabitants and with buildings separated by less than 200 m; otherwise the commune was defined as rural (Insee 2016). Assignment of the rural or urban status was based on the census the closest in time to each address period. INSEE provided 11 population censuses available from 1926 to 2006.

The rural or urban status was determined for 8,077 residential addresses geocoded. The status was missing for 778 (8.8%) addresses because they were not geocoded (no address or incomplete address), they were located abroad or the starting or ending year of the residency period was missing so no census could be allocated to the address. An address could be divided into several periods corresponding to a census and a status. A total of 1,814 (22.5%) addresses were defined as rural and 6,496 (80.4%) as urban. For 233 (2.9%) addresses, the status rural or urban has changed between 1925 and 2014.

Among occupational addresses geocoded, the rural or urban status was assigned to 7,732 of them and was missing for 804 (9.4%) addresses, because they were not geocoded, located abroad and not allocated a census. There were 1,023 (13.2%) rural and 6,775 (87.6%) urban occupational addresses. A change of status occurred for 66 (0.9%) occupational addresses between 1925 and 2014.

In fine, the urban or rural status of the residence and residency periods were used to compute the number of years of living in an urban area from birth to inclusion in the E3N cohort that was considered as a surrogate for airborne dioxin exposure to account for past exposures in the statistical analysis.

Appendice 6. Publication by Binachon et al. 2014

Life in urban areas and breast cancer risk in the French E3N cohort

Blandine Binachon¹, Laure Dossus^{1,2,3}, Aurélie MN Danjou^{1,4}, Françoise Clavel-Chapelon^{2,3,5} and Béatrice Fervers^{1,4}

¹ Unité Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France

² Inserm, Centre for Research in Epidemiology and Population Health (CESP), U1018, Nutrition, Hormones and Women's Health Team, 94805 Villejuif, France

³ Université Paris Sud, UMRS 1018, 94805 Villejuif, France

⁴ Laboratoire Santé, Individu, Société, Université de Lyon, Lyon, France

⁵ IGR, 94805 Villejuif, France

Reference:

Binachon B, Dossus L, Danjou AMN, Clavel-Chapelon F, Fervers B. 2014. Life in urban areas and breast cancer risk in the French E3N cohort. *European Journal of Epidemiology* 29:743–751; doi:10.1007/s10654-014-9942-z.

Author's final draft post-refereeing

Life in urban areas and breast cancer risk in the French E3N cohort

B. Binachon^a, L. Dossus^{a,b,c}, A.M.N. Danjou^{a,d}, F. Clavel-Chapelon^{b,c,e}, B. Fervers^{a,d*}

^a Centre Léon Bérard, Unité Cancer et Environnement, 28 rue Laënnec, 69373 Lyon cedex 08, France.

^b Inserm, Centre for research in Epidemiology and Population Health (CESP), U1018, Nutrition, Hormones and Women's Health team, F-94805, Villejuif, France.

^c Université Paris Sud, UMRS 1018, F-94805, Villejuif, France.

^d Université de Lyon, laboratoire Santé, Individu, Société, Lyon, France.

^e IGR, F-94805, Villejuif, France.

***Corresponding author at:** Unité Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon cedex 08, France. Tel.: +33478782801. Fax: +33478782883. Email: Beatrice.fervers@lyon.unicancer.fr.

Conflict of interest: none

Published in European Journal of Epidemiology (2014)

Author's final draft post-refereeing

Abstract (words: 200)

Background

It has been suggested that women living in urban areas are at higher risk of breast cancer (BC) compared to women living in rural areas. However, most published studies on this topic are ecological and did not adjust for individual BC risk factors. We investigated this hypothesis in a French prospective cohort.

Methods

Cox proportional hazards regression models were used to evaluate the association between birth or residence in an urban area and BC risk among 75,889 women of the French E3N cohort (aged 38 to 66 years at recruitment in 1990) before and after adjustment for known BC risk factors and stratifying on birth cohort.

Results

From 1990 to 2008, a total of 5,145 BC cases were diagnosed. Being born in an urban area was associated with BC risk before (HR: 1.11, 95% CI: 1.05-1.18) and after (HR: 1.07, 95% CI: 1.01-1.14) adjustment for known BC risk factors. Living in an urban area in 1990 was not associated with BC risk.

Conclusion

Being born in an urban area is associated with a weak increase in BC risk. This may be suggestive of higher exposure to air pollution and to other environmental exposures, to be investigated in future studies.

Author's final draft post-refereeing

Introduction

Breast cancer is the most frequent cancer in women worldwide. Incidence rates have been increasing in the last decades with highest incidence in developed countries [1,2]. A higher breast cancer incidence is generally observed in large cities than in rural areas [1,3]. Since 1900, urbanisation and rural flight resulted, in France and other developed countries, in a growing number of urban areas at the expense of rural ones [4]. A number of studies, conducted mostly in developed countries, compared breast cancer risk among women living in urban and rural areas. All of them were ecological and used population-based registries [5–10], except for one case-control study [9]. Most of these studies observed an increased breast cancer risk among women living in urban areas compared to those living in rural areas. The average increase in risk was 15%. Only one study, conducted in Australia, did not find any statistically significant difference in breast cancer incidence between rural and urban areas [10]. Although some known breast cancer risk factors, such as reproductive and lifestyle factors, vary between urban and rural populations, no adjustment was made for breast cancer individual risk factors, except for one study [9].

To our knowledge, no prospective epidemiologic study has provided results about the association between residence in an urban area and breast cancer risk. We used data from the large French cohort study E3N to investigate the association between urban or rural status of the birthplace and residence during adulthood, and breast cancer risk, adjusting for most known individual breast cancer risk factors.

Author's final draft post-refereeing

Materials and Methods

The E3N cohort study

The methodology and design of the E3N (Etude Epidémiologique auprès des femmes de la Mutuelle Générale de l'Education Nationale) study has been detailed elsewhere [11]. Briefly, E3N is a prospective cohort initiated in 1990 and including 98,995 French women affiliated to the Mutuelle Générale de l'Education Nationale (MGEN), a national health insurance plan covering primarily schoolteachers. Participants were 38 to 66 years old when they entered the cohort in 1990. They provided written informed consent for external health follow-up through the health insurer and completed biennial self-administered questionnaires on established risk factors for breast cancer, including aspects of reproductive life, menopausal status, history of benign breast disease, breast cancer in first-degree relatives and lifestyle. E3N was approved by the French Commission for Data Protection and Privacy.

Study population

Women with prevalent cancer at inclusion (N=4,840), lack of follow-up (N=2,085), unknown date of breast cancer diagnosis (N=78) as well as women with phyllode tumors (N=6) were excluded from the present study. We further excluded women who did not provide information on place of birth, those born out of metropolitan France, as well as women with miscoded postal codes or addresses for place of birth and/or residence at study entry (N=16,097).

Breast cancer case ascertainment

Occurrence of cancer was self-reported in each E3N biennial questionnaire. A small number of cases were further identified from the insurance files or information on causes of death

Author's final draft post-refereeing

obtained from the National Service on Causes of Death. Pathology reports were used to confirm the diagnosis of invasive breast cancer and were obtained for 93% of declared incident invasive breast cancer cases. Because the proportion of false-positive self-reports was low (<5%), we also included participants who reported a breast cancer diagnosis with no pathology report available.

Definition of urban and rural areas

E3N women's residence in urban area was assessed at two different times: at the time of birth and at the time of their recruitment into the cohort in 1990. The place of birth (postal code and commune) was obtained in the first self-administered questionnaire. The commune and postal code of participants residence in 1990 was provided by the MGEN. In France, the commune is the smallest administrative division of the territory. According to the National Institute of Geographic Information (IGN) and the National Research Institute for Demography and Population (INED), an urban area was defined as a commune or a group of communes containing more than 2,000 inhabitants aggregated (i.e. with buildings not separated by more than 200 meters); other areas were defined as rural [12–14]. Classification of urban/rural status of place of birth was based on data from the closest census: the 1926 census was used to attribute the urban/rural status of birthplace of women born in 1924 to 1928, the census of 1931 for women born in 1929 to 1933, the census of 1936 for women born in 1934 to 1941, the census of 1946 for women born in 1942 to 1949 and the census of 1954 for women born in 1950 and 1951. The urban/rural status of the place of residence in 1990 was based on the 1990 census. Urban areas were further stratified according to the number of inhabitants, with the following thresholds: 5,000; 10,000; 20,000; 50,000; 100,000; and 500,000 inhabitants [15].

Author's final draft post-refereeing

Statistical analysis

Person-years for the present analysis were calculated from the date of return of the first questionnaire (1990) to the date of diagnosis of a first primary invasive breast cancer, the date of the last questionnaire completed, the date of death, or June 2008 (date of end of follow-up for the present study), whichever occurred first.

Cox's proportional hazards regression models stratified by 5-year-interval birth cohorts with the women's age as the time scale were used to estimate HR and 95% Confidence Intervals (CI) for breast cancer according to status of the places of birth and residence in 1990. Participants born/living in an urban area were compared with participants born/living in a rural area (reference class). The HRs of breast cancer were calculated overall and stratified by birth cohort (<1930; 1930-1934; 1935-1939; 1940-1944; \geq 1945). In multivariable analyses, adjustment was made *a priori* for several known breast cancer risk factors, including height (continuous, cm), Body Mass Index (BMI, WHO cut points: 20, 25, 30 kg/m²), physical activity (quartiles, cut points: 34, 47, 62 Mets-h/week), energy intake without alcohol (quartiles, cut points: 1735, 2078, 2468 kcal/day), alcohol intake (no intake, <6.9 and \geq 6.9 g/day), smoking status (never smoker, current smoker and former smoker), education level (undergraduate, graduate with 1 to 4 years university degree and graduate with 5 or more years university degree), age at menarche (cut points: 12 and 14 years), previous use of oral contraceptives (ever/never), use of oral progestagens alone before menopause (ever/never), parity (number of full-term pregnancies: 0, 1, 2 and 3 or more), age at first full-term pregnancy (before and after 30 years old), breastfeeding (yes/no), menopausal status (yes/no), age at menopause (cut points: 47 and 54 years old among postmenopausal women only), mammographic exam during the previous follow-up period (yes/no), use of menopausal hormone therapy (ever, never and premenopausal women), previous family history of breast cancer (yes/no), previous personal history of benign breast disease (ever/never). BMI,

Author's final draft post-refereeing

smoking status, menopausal status, mammographic exam during the previous follow-up period, use of menopausal hormone therapy, and previous personal history of benign breast disease were included in the models as time-dependent variables. For variables considered as time-dependent, data recorded in questionnaires n and earlier were used to prospectively categorize women for the period that followed (i.e., between questionnaire n and questionnaire $n+1$, where $n+1$ was the next completed questionnaire). When fewer than 5% of the values of a covariate were missing, they were replaced with the mode or the median values observed among the subjects with complete data. When there were more than 5% missing values we treated them as a separate category. Trend tests across categories of number of inhabitants in place of birth/residence were performed using quantitative scores of 1 (rural area), 2, 3, 4, 5 and treating them as continuous variables. Interactions were tested with all potential confounders. All tests of statistical significance were two-sided and the level of significance considered was 0.05. The SAS statistical software (SAS Institute Inc., Gary, North Carolina, version 9.3) was used for statistical analysis.

Results

Among the 75,889 women eligible for the current study, a total of 5,145 incident primary invasive breast cancer cases were diagnosed during 1,176,903 person-years of follow-up (mean duration: 15.5 years, SD 4.5 years). The mean age was 49.2 years (SD: 6.6) at recruitment and 59.2 years (SD: 7.7) at breast cancer diagnosis. Overall 69% percent of the women were born in an urban area and 67% lived in an urban area in 1990. These percentages varied across birth cohorts as 63% of the women were born in urban areas before 1930 while they were up to 72% after 1945.

Author's final draft post-refereeing

Baseline characteristics of the participants according to rural vs. urban status of their place of birth or residence in 1990 are provided in Table 1. Compared to women living or born in rural areas, those living or born in urban areas had slightly less recreational physical activity, a higher alcohol consumption, were more frequently current or ex-smokers, had a higher education level, a lower age at menarche, they had more often no child or only one child and had their first full-term pregnancy later, they had more often breastfed at least one child, and they had more frequently had a previous mammography.

HRs of breast cancer according to urban versus rural status are presented in Table 2. Compared to women born in rural areas, women born in urban areas were at significantly increased risk of breast cancer before and after adjustment on potential confounders (HR: 1.11, 95% CI: 1.05-1.18 and HR: 1.07, 95% CI: 1.01-1.14, respectively). When further stratifying on urban birthplaces according to number of inhabitants, statistically significant trends were observed with the highest HR for women born in communes with 50,000 to 99,999 inhabitants and 100,000 to 499,999 inhabitants (adjusted HR: 1.15, 95% CI: 1.03-1.27 and HR: 1.15, 95% CI: 1.04-1.28, respectively).

No significant association was observed between the urban/rural status of the place of residence in 1990 and breast cancer risk. The statistically significant association observed for women living in communes of more than 500,000 inhabitants compared to those living in rural areas (HR: 1.13, 95% CI: 1.00-1.27) was no longer present after adjustment for individual risk factors (HR: 0.98, 95% CI: 0.87-1.11). Women that were born and lived in 1990 in an urban area were not at increased risk of breast cancer compared to those born and living in rural areas (HR: 1.02, 95% CI: 0.94-1.12).

There was no statistically significant interaction with any of the potential confounding factors. Restricting the analyses to women with a previous mammography (women who had at least one mammography during their follow-up, between 1990 and 2008, 4,003 cases) did not alter our

Author's final draft post-refereeing

results (data not shown). Stratification by birth cohort yielded statistically significant association among women born between 1935 and 1944 (figure 1). An increasing risk of breast cancer was observed with increasing number of inhabitants ($P_{\text{trend}}=0.04$ and 0.02 for women born in 1935-1939 and 1940-1944).

Discussion

This is the first prospective study investigating the association between birth and residence in urban versus rural areas and risk of breast cancer while adjusting for a large number of potential confounders. We found a positive association between birth in an urban area and breast cancer risk with a positive trend for increasing number of inhabitants, although the overall risk was small and no consistent dose-response was observed with respect to the size of the urban area where each subject was born. There was no association between living in an urban area in adulthood and breast cancer risk.

Only few retrospective ecological studies previously evaluated the association between breast cancer risk and place of residence [5–10]. In these studies, the timing of living in urban settings was the study entry and none of them considered the birthplace. Except for one Australian study that was not conclusive [10], all studies concluded to a higher breast cancer incidence among women living in urban areas. Only one case-control study adjusted for individual breast cancer risk factors, such as menopausal status, benign breast cancer disease history and lifestyle factors [9].

Socioeconomic status is associated with breast cancer risk and it is likely to be different between rural and urban residents [16]. The E3N cohort is very homogeneous in terms of socio-economic status as it is composed mostly of schoolteachers. However, we did observe that urban residents had more often a higher university degree than rural residents and this

Author's final draft post-refereeing

was controlled for in our multivariable models. When comparing lifestyle and reproductive characteristics between urban and rural residents of the E3N cohort, women born and living in urban areas seemed to have less healthy lifestyles than rural residents: they had less recreational activity and a higher proportion of them smoked and drunk alcohol regularly, three factors that are well known breast cancer risk factors [17] [18]. Furthermore, urban residents seemed to have higher endogenous and exogenous hormonal exposures (also known to increase breast cancer risk [19]), compared to rural residents, as they had earlier menarche, used more often hormonal therapies (oral progestagens alone before menopause and menopausal hormone therapy), had a lower number of children, and tended to have their first child later. Although the effects of these factors are controlled for in our multivariable analyses, we cannot exclude that some residual confounding may still be present and affect the results. Dietary patterns and food consumption may differ between urban and rural residents [20] and dietary patterns have been shown to be related to breast cancer risk [21,22]. In our population, dietary patterns were very similar between urban and rural residents and adjustment for these patterns did not affect risk estimates (data not shown). Dietary intakes at birth or in childhood, however, were not available. Urban residents may have lower level of sunlight exposure and suffer more often from vitamin D deficiency than rural residents [23] and low vitamin D concentrations were found associated with an increased breast cancer risk [24,25]. However, data available in the E3N cohort did not allow us to adjust for sun exposure.

The main strengths of this study were its prospective design, the small number of women lost to follow-up and the regular update of breast cancer risk factors during follow-up. Adjustment for various potential confounders decreased the possibility that the HRs observed were explained by confounding. The number of inhabitants per commune was obtained from the

Author's final draft post-refereeing

official census, performed regularly every five years (except during World War II: 10 years).

The precision is therefore high. Urban and rural status of each location were assigned by geographers based on the number of inhabitants aggregated, complying with the official definition of the French National Geographic Institute for each period [26].

Nevertheless, the approach used in our study has some limitations. First, geographical boundaries between urban and rural areas have evolved in France since the beginning of the last century. Although the threshold of 2,000 inhabitants has been criticized [26], there is currently no consensus in the definition of thresholds for urban and rural status and a large variation exists among European countries, from 200 inhabitants in Denmark to 10,000 inhabitants in Greece [14]. Furthermore, in France the number of people living in urban areas has been growing during the last century: from 50% in 1936 to two-thirds in 1990 [27]. This is due to a high proportion of rural residents moving to the cities but also to the extension and industrialization of the cities. The boundaries between rural and urban areas became less and less clear while some “peri-urban” intermediate areas appeared. However we choose to keep the threshold of 2,000 inhabitants to separate urban from rural areas to allow comparison between studies [28]. At the time the participants of our study were born, i.e. between 1924 and 1951, the boundaries between urban and rural areas were more precise than in 1990. This may explain the absence of association between living in an urban area in 1990 and breast cancer risk. We performed additional analyses with different population thresholds to go beyond the traditional urban/rural dichotomy and take into account the variations in the definitions of the urbanicity [29]. In these analyses, however, no clear trend was observed in the association with increasing population size.

Furthermore, the 1990 E3N questionnaire asked for the women's place of birth. However, in France, the place of birth may be the location of the maternity and thus different from the woman's place of residence at birth, ie the parents' residence at the time of her birth. From

Author's final draft post-refereeing

1930, an increasing number of women in France gave birth in hospitals [30]. While 80% of births occurred at home in 1931, 53% of the deliveries occurred in hospitals in 1952 [31]. It is difficult to obtain precise data about this trend but it is likely to result in a misclassification bias in our study (classification of women living in rural areas into the urban category) which, although non-differential between cases and non-cases, might have diluted our estimates, in particular for the most recent birth cohorts. After stratification by birth cohort, we found a positive association only for the women born between 1935 and 1944, but no association for the women born before 1935 or after 1944. Before 1935 this may be explained by a lower level of air pollution, and after 1944 by a higher risk of misclassification bias due to the increasing proportion of women giving birth in hospitals.

Birth and residence in urban places may be suggestive of higher exposure to air pollution, released from road traffic, industrial facilities, use of fossil fuels for power generation and waste incineration [32]. Association of exposure to air pollution (in particular in urban areas) with a decrease in life expectancy in general [33–35] and an increase in non-accidental, cardiovascular, ischemic heart diseases [36], and lung-cancer mortality [37,38], is well established. More recently, several studies have shown an association between breast cancer incidence and air pollution [39,40], in particular for early life periods [41,42]. Furthermore, breast cancer has been associated with environmental exposure to several carcinogens with endocrine disrupting properties, such as PCBs (polychlorinated biphenyls) [43] or dioxins [44,45].

Our finding of an association between being born in an urban area and breast cancer risk and an absence of association with the residence in an urban area at adulthood is suggestive of a deleterious effect of exposure in early life. It has been suggested previously that the impact of early life exposures (i.e. in utero and during childhood) might be more important compared to later life exposure [46] due to greater sensitivity of breast tissue during early life [47].

Author's final draft post-refereeing

In conclusion, our findings suggest that being born in an urban area increases the risk of breast cancer. In future studies, it would be interesting to consider the entire residential history of the women as well as their workplace history. Further prospective studies should also investigate the associations between air pollution, and other environmental factors, and breast cancer risk.

Author's final draft post-refereeing

Aknowledgments

We thank all participants for providing the data used in this study and practitioners for providing pathology reports. We are grateful to Alice Vilier and Anne Bijon for managing the data. We thank Pr Denise Pumain (Université Paris I), Liliane Lizzi (Unité Mixte de Recherche Géographie-Cités, Paris, France), Anne Bretagnolle (Unité Mixte de Recherche Géographie-Cités, Paris, France), and Benoît Riandey (Institut National d'Etudes Démographiques, France) who kindly provided geographical data.

Author's final draft post-refereeing

Reference List

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 127: 2893-2917. 10.1002/ijc.25516 [doi].
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D (2011) Global cancer statistics. *CA Cancer J Clin* 61: 69-90. caac.20107 [pii];10.3322/caac.20107 [doi].
3. Canadian Breast Cancer Initiative, Health Canada (2001) Review of Lifestyle and Environmental Risk Factors for Breast Cancer - Summary Report.
4. Julien P (1998) L'urbanisation se poursuit dans les années 90.
5. Rougereau A, Pottier D, Robillard J, Mace-Lesech J, Henry-Amar M, Launoy G (1993) Incidence du cancer colo-rectal et du cancer du sein dans le département du Calvados. 24/1993.
6. Minelli L, Stracci F, Casseti T, Canosa A, Scheibel M, Sapia IE, Romagnoli C, La RF (2007) Urban-rural differences in gynaecological cancer occurrence in a central region of Italy: 1978-1982 and 1998-2002. *Eur J Gynaecol Oncol* 28: 468-472.
7. Krzyzak M, Maslach D, Juczewska M, Lasota W, Rabczenko D, Marcinkowski J, Szpak A (2010) Differences in breast cancer incidence and stage distribution between urban and rural female population in Podlaskie Voivodship, Poland in years 2001-2002. *Ann Agric Environ Med* 17: 159-162. 17159 [pii].
8. Dey S, Soliman AS, Hablas A, Seifeldein IA, Ismail K, Ramadan M, El-Hamzawy H, Wilson ML, Banerjee M, Boffetta P, Harford J, Merajver SD (2010) Urban-rural differences in breast cancer incidence in Egypt (1999-2006). *Breast* 19: 417-423. S0960-9776(10)00112-8 [pii];10.1016/j.breast.2010.04.005 [doi].
9. Robert SA, Strombom I, Trentham-Dietz A, Hampton JM, McElroy JA, Newcomb PA, Remington PL (2004) Socioeconomic risk factors for breast cancer: distinguishing individual- and community-level effects. *Epidemiology* 15: 442-450. 00001648-200407000-00011 [pii].
10. Wilkinson D, Cameron K (2004) Cancer and cancer risk in South Australia: what evidence for a rural-urban health differential? *Aust J Rural Health* 12: 61-66. 10.1111/j.1038-5282.2004.00555.x [doi];AJR555 [pii].
11. Clavel-Chapelon F, van Liere MJ, Giubout C, Niravong MY, Goulard H, Le CC, Hoang LA, Amoyel J, Auquier A, Duquesnel E (1997) E3N, a French cohort study on cancer risk factors. E3N Group. Etude Epidemiologique aupres de femmes de l'Education Nationale. *Eur J Cancer Prev* 6: 473-478.
12. Institut Géographique National. (2013)
13. Institut National d'Etudes Démographiques. (2013)
14. Houillon V, Thomsin L (2001) Définitions du rural et de l'urbain dans quelques pays européens. *Espaces, Population, Société* 2001-1-2.

Author's final draft post-refereeing

15. United Nations Statistics Division (2013) Demographic and Social Statistics, population density and urbanization.
16. Newman LA, Martin IK (2007) Disparities in breast cancer. *Curr Probl Cancer* 31: 134-156. S0147-0272(07)00004-9 [pii];10.1016/j.currprobcancer.2007.01.003 [doi].
17. Cui Y, Miller AB, Rohan TE (2006) Cigarette smoking and breast cancer risk: update of a prospective cohort study. *Breast Cancer Res Treat* 100: 293-299. 10.1007/s10549-006-9255-3 [doi].
18. Scoccianti C, Straif K, Romieu I (2013) Recent evidence on alcohol and cancer epidemiology. *Future Oncol* 9: 1315-1322. 10.2217/fon.13.94 [doi].
19. Adami, H. O., Hunter, D., and Trichopoulos, D. (2008) *Textbook of Cancer Epidemiology*, Second Edition . Oxford University Press.
20. Wang CN, Liang Z, Wei P, Liu P, Yu JX, Zhang DM, Ma FL (2002) Changes in dietary patterns and certain nutrition-related diseases in urban and rural residents of Jiangsu Province, China, during the 1990s. *Biomed Environ Sci* 15: 271-276.
21. World Cancer Research Fund, American Institute for Cancer Research (2007) *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective*.
22. Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F, Boutron-Ruault MC (2009) Postmenopausal breast cancer risk and dietary patterns in the E3N-EPIC prospective cohort study. *Am J Epidemiol* 170: 1257-1267. kwp257 [pii];10.1093/aje/kwp257 [doi].
23. Choi EY (2012) 25(OH)D status and demographic and lifestyle determinants of 25(OH)D among Korean adults. *Asia Pac J Clin Nutr* 21: 526-535.
24. Engel P, Fagherazzi G, Mesrine S, Boutron-Ruault MC, Clavel-Chapelon F (2011) Joint effects of dietary vitamin D and sun exposure on breast cancer risk: results from the French E3N cohort. *Cancer Epidemiol Biomarkers Prev* 20: 187-198. 1055-9965.EPI-10-1039 [pii];10.1158/1055-9965.EPI-10-1039 [doi].
25. Shao T, Klein P, Grossbard ML (2012) Vitamin D and breast cancer. *Oncologist* 17: 36-45. theoncologist.2011-0278 [pii];10.1634/theoncologist.2011-0278 [doi].
26. Pumain D, Riandey B (1986) *Le Fichier de l'Ined : "urbanisation de la France"*. Espace, populations, sociétés .
27. Fanouillet J-C, Madinier C (1996) L'extension des villes de 1936 à 1990. *INSEE PREMIERE* 451.
28. Francis R, Chadwick M (2013) *Urban Ecosystem: Understanding the Human Environment*.
29. Hall SA, Kaufman JS, Ricketts TC (2006) Defining urban and rural areas in U.S. epidemiologic studies. *J Urban Health* 83: 162-175. 10.1007/s11524-005-9016-3 [doi].
30. Morel M-F (2010) *Naître en France du XVII au XXème siècles*. Contact Santé .

Author's final draft post-refereeing

31. Morel M-F (2007) Histoire de la naissance en France (XVIIe–XXe siècle). ADSP .
32. United States Environmental Protection Agency (2001) U.S. Toxics Release Inventory Public Data Release.
33. Pope CA, III, Ezzati M, Dockery DW (2009) Fine-particulate air pollution and life expectancy in the United States. *N Engl J Med* 360: 376-386. 360/4/376 [pii];10.1056/NEJMsa0805646 [doi].
34. Correia AW, Pope CA, III, Dockery DW, Wang Y, Ezzati M, Dominici F (2013) Effect of air pollution control on life expectancy in the United States: an analysis of 545 U.S. counties for the period from 2000 to 2007. *Epidemiology* 24: 23-31. 10.1097/EDE.0b013e3182770237 [doi].
35. Crouse DL, Goldberg MS, Ross NA, Chen H, Labreche F (2010) Postmenopausal breast cancer is associated with exposure to traffic-related air pollution in Montreal, Canada: a case-control study. *Environ Health Perspect* 118: 1578-1583. 10.1289/ehp.1002221 [doi].
36. Cesaroni G, Badaloni C, Gariazzo C, Stafoggia M, Sozzi R, Davoli M, Forastiere F (2013) Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ Health Perspect* 121: 324-331. 10.1289/ehp.1205862 [doi].
37. Loomis D, Grosse Y, Lauby-Secretan B, El GF, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K, on behalf of the International Agency for Research on Cancer Monograph Working Group IARC (2013) The carcinogenicity of outdoor air pollution. *Lancet Oncol* 14: 1262-1263.
38. Nyberg F, Gustavsson P, Jarup L, Bellander T, Berglind N, Jakobsson R, Pershagen G (2000) Urban air pollution and lung cancer in Stockholm. *Epidemiology* 11: 487-495.
39. Chen F, Bina WF (2012) Correlation of white female breast cancer incidence trends with nitrogen dioxide emission levels and motor vehicle density patterns. *Breast Cancer Res Treat* 132: 327-333. 10.1007/s10549-011-1861-z [doi].
40. Wei Y, Davis J, Bina WF (2012) Ambient air pollution is associated with the increased incidence of breast cancer in US. *Int J Environ Health Res* 22: 12-21. 938209397 [pii];10.1080/09603123.2011.588321 [doi].
41. Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Muti P, Trevisan M, Edge SB, Freudenheim JL (2005) Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. *Cancer Epidemiol Biomarkers Prev* 14: 53-60. 14/1/53 [pii].
42. Nie J, Beyea J, Bonner MR, Han D, Vena JE, Rogerson P, Vito D, Muti P, Trevisan M, Edge SB, Freudenheim JL (2007) Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. *Cancer Causes Control* 18: 947-955. 10.1007/s10552-007-9036-2 [doi].
43. Lauby-Secretan B, Loomis D, Grosse Y, El GF, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K (2013) Carcinogenicity of polychlorinated biphenyls and polybrominated biphenyls. *Lancet Oncol* 14: 287-288. S1470-2045(13)70104-9 [pii];10.1016/S1470-2045(13)70104-9 [doi].

Author's final draft post-refereeing

44. Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA (2007) Environmental pollutants and breast cancer: epidemiologic studies. *Cancer* 109: 2667-2711. 10.1002/cncr.22655 [doi].
45. Warner M, Eskenazi B, Mocarelli P, Gerthoux PM, Samuels S, Needham L, Patterson D, Brambilla P (2002) Serum dioxin concentrations and breast cancer risk in the Seveso Women's Health Study. *Environ Health Perspect* 110: 625-628. sc271_5_1835 [pii].
46. Potischman N, Troisi R (1999) In-utero and early life exposures in relation to risk of breast cancer. *Cancer Causes Control* 10: 561-573.
47. Russo J (2004) Molecular basis of breast cancer.

Table 1. Main baseline characteristics of the study population by urban/rural status of the places of birth and residence in 1990, E3N cohort, France, 1990-2008 (N=75,889)

Baseline characteristics (% or mean (SD))	Missing data (%)	Place of birth		Place of residence in 1990	
		Rural (N=23,705)	Urban (N=52,184)	Rural (N=25,243)	Urban (N=50,646)
Age (years)	0	49.9 (6.7)	48.9 (6.6)	48.9 (6.6)	49.4 (6.6)
Height (cm)	<0.001	161.3 (5.6)	161.9 (5.7)	161.7 (5.6)	161.8 (5.7)
BMI (kg/m²)	1.9				
<20		17.7	19.0	17.5	19.2
20-25		64.7	63.8	63.7	64.3
25-30		14.6	14.1	15.3	13.7
>30		3.0	3.1	3.5	2.9
Recreational physical activity (METs-h/wk)	0				
<34		43.6	47.1	40.7	48.6
34-47		23.7	22.3	23.5	22.4
47-62		14.9	14.4	15.8	14.0
>62		17.8	16.2	20.0	15.1
Total energy intake without alcohol (kcal/day)	21.7				
<1735		23.8	25.0	23.6	25.0
1735-2078		25.3	25.1	25.2	25.1
2078-2468		25.5	25.0	25.2	25.1
>2468		25.4	25.2	26.1	24.8
Alcohol intake (g/day)	21.7				
None		13.4	12.0	13.2	12.0
<6.9		40.1	37.0	40.0	37.0
≥6.9		46.5	51.0	46.9	51.0
Smoking status	0.3				
Never smoker		58.8	52.1	56.9	52.9
Ex-smoker		29.1	32.1	30.1	31.6
Current smoker		12.1	15.9	13.0	15.6
Education level	3.9				

Author's final draft post-refereeing

Baseline characteristics (% or mean (SD))	Missing data (%)	Place of birth		Place of residence in 1990	
		Rural (N=23,705)	Urban (N=52,184)	Rural (N=25,243)	Urban (N=50,646)
Undergraduate		16.7	12.4	14.9	13.1
1 to 4 years university degree		71.8	67.5	74.7	66.0
5 or more years university degree		11.5	20.1	10.4	20.9
Age at menarche	2.1				
<12		18.0	21.9	19.8	21.1
12-14		49.7	50.8	50.0	50.7
>14		32.4	27.3	30.2	28.3
Ever use of oral contraceptive	0	56.9	62.7	60.4	61.1
Ever use of oral progestagens alone before menopause	0	44.1	47.1	44.9	46.8
Postmenopausal at recruitment	5.0	44.9	39.8	39.4	42.5
Parity	0.8				
No child		10.5	12.6	9.3	13.3
1 child		16.7	16.4	15.5	17.0
2 children		44.1	42.3	44.1	42.2
3 or more children		28.6	28.8	31.1	27.6
Age at first full-term pregnancy (y)	12.9				
<30		89.7	86.9	89.9	86.7
≥30		10.3	13.1	10.1	13.3
Breastfeeding (ever)	22.6	40.9	47.6	43.3	46.7
Age at menopause (years)	5.0				
<47		12.0	12.2	12.0	12.2
47-54		72.1	72.8	72.4	72.6
>54		16.0	15.2	15.6	15.2
Previous mammography	0	70.6	71.7	66.9	73.6
Ever use of menopausal hormone therapy	13.6	21.7	20.3	18.4	21.9
Previous family history of breast cancer	0.8	11.0	11.6	11.1	11.6
Personal history of benign breast disease	0	22.3	22.4	21.1	22.9

Author's final draft post-refereeing

Table 2. HRs of breast cancer according to urban or rural status and number of inhabitants of the place of birth and the place of residence in 1990, E3N cohort, France, 1990-2008 (N=75,889)

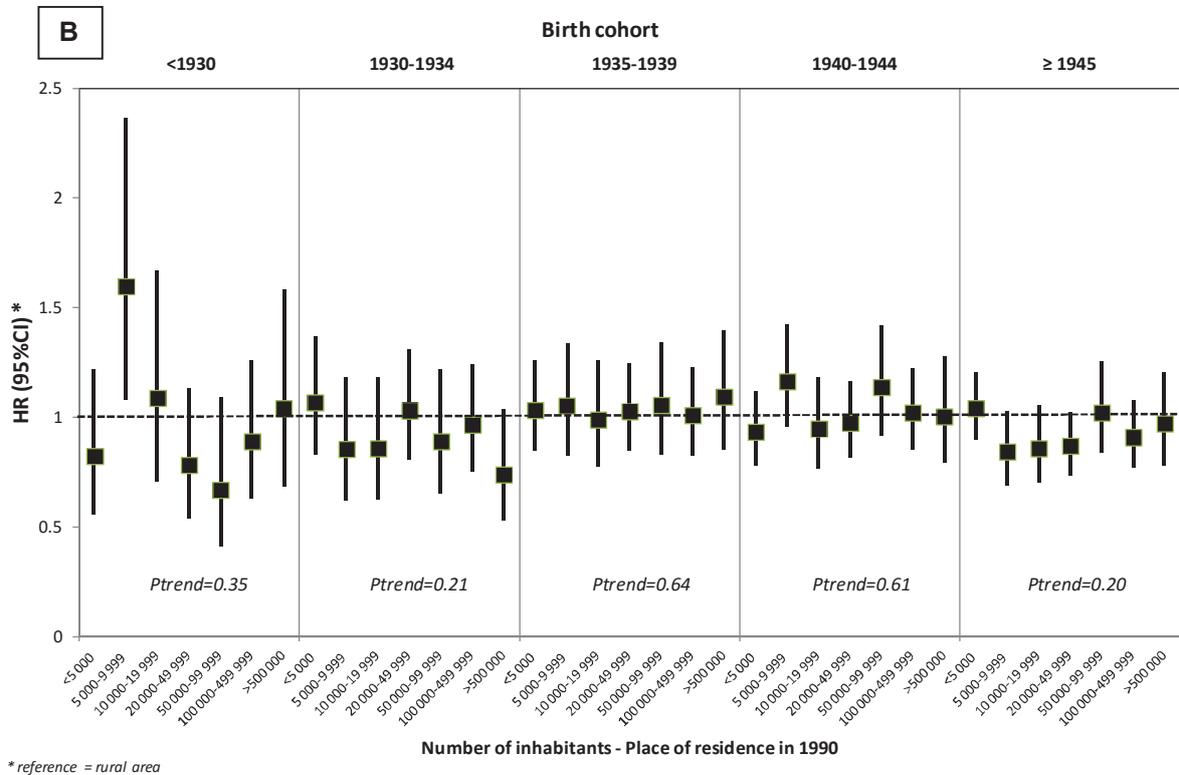
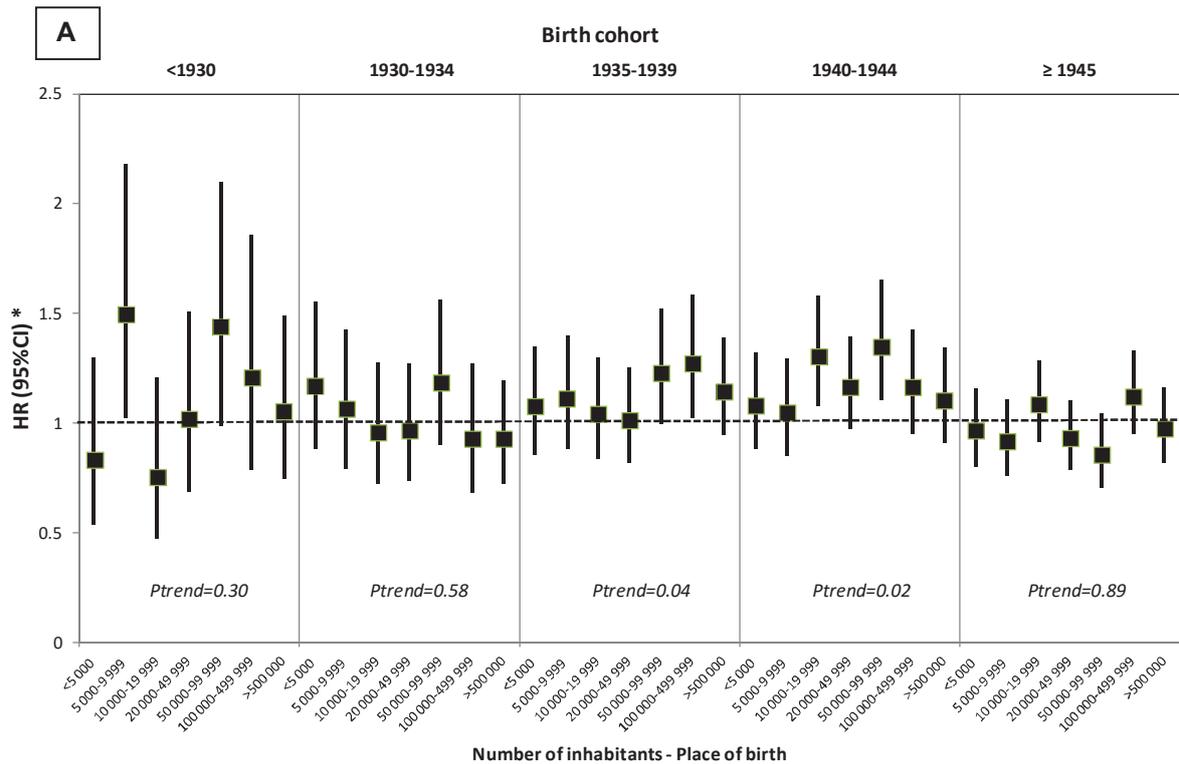
	Place of birth			Place of residence in 1990			
	Person-Years	Cases	HR (95% CI) ^a	Person-Years	Cases	HR (95% CI) ^a	HR (95% CI) ^b
Rural area	369,359	1,518	1.00 (reference)	393,777	1,666	1.00 (reference)	1.00 (reference)
Urban area	807,543	3,627	1.11 (1.05-1.18)	783,125	3,479	1.04 (0.98-1.10)	0.98 (0.92-1.04)
Less than 5,000 inhabitants	110,228	469	1.06 (0.95-1.17)	156,998	684	1.03 (0.94-1.12)	1.01 (0.92-1.10)
5,000 to 9,999 inhabitants	104,351	446	1.06 (0.96-1.18)	90,091	404	1.05 (0.95-1.18)	1.01 (0.91-1.13)
10,000 to 19,999 inhabitants	117,690	538	1.13 (1.03-1.25)	87,971	366	0.98 (0.87-1.09)	0.93 (0.83-1.04)
20,000 to 49,999 inhabitants	139,477	592	1.06 (0.96-1.16)	153,948	666	1.01 (0.93-1.11)	0.95 (0.87-1.04)
50,000 to 99,999 inhabitants	96,680	466	1.19 (1.07-1.32)	80,167	372	1.08 (0.97-1.21)	1.02 (0.91-1.14)
100,000 to 499,999 inhabitants	105,056	505	1.20 (1.09-1.33)	141,994	640	1.05 (0.96-1.15)	0.97 (0.89-1.07)
More than 500,000 inhabitants	134,061	611	1.12 (1.02-1.23)	71,957	347	1.13 (1.00-1.27)	0.98 (0.87-1.11)
P value for trend			0.0001			0.06	0.38

^a Stratified by birth cohorts and age-adjusted

^b Stratified by birth cohorts and adjusted for age, height (cm), Body Mass Index (BMI, cut points: 20, 25, 30 kg/m²), physical activity in 1990 (cut points: 34, 47, 62 Met-h/week), energy intake without alcohol (cut points: 1735, 2078, 2468 kcal/day), alcohol intake (no intake, <6.9 and ≥6.9 g/day), smoking status (never smoker, current-smoker and former smoker), total education level (undergraduate, graduate with 1 to 4 years university degree and graduate with 5 or more years university degree), age at menarche (cut points: 12 and 14 years), previous use of oral contraceptives (ever/never), use of oral progestagens alone before menopause (ever/never), parity (number of full-term pregnancies: 0, 1, 2 and 3 or more), age at first full-term pregnancy (before and after 30 years old), breastfeeding (yes/no), menopausal status (yes/no), age at menopause (cut points: 47 and 54 years old among postmenopausal women only), mammographic exam during the previous follow-up period (yes/no), use of menopausal hormone therapy (ever, never and premenopausal women), previous family history of breast cancer (yes/no), previous personal history of benign breast disease (ever/never)

Author's final draft post-refereeing

Figure1. HRs of breast cancer according to number of inhabitants of the place of birth (A) and the place of residence in 1990, stratified by birth cohort, E3N cohort, France, 1990-2008 (N=75,889)



Adjusted for height (cm), Body Mass Index (BMI, cut points: 20, 25, 30 kg/m²), physical activity in 1990 (cut points: 34, 47, 62 Met-h/week), energy intake without alcohol (cut points: 1735, 2078, 2468 kcal/day), alcohol intake (no intake, <6.9 and ≥6.9 g/day), smoking status (never smoker, current-smoker and former smoker), total education level (undergraduate, graduate with 1 to 4 years university degree and graduate with 5 or more years university degree), age at menarche (cut points: 12 and 14 years), previous use of oral contraceptives (ever/never), use of oral progestagens alone before menopause (ever/never), parity (number of full-term pregnancies: 0, 1, 2 and 3 or more), age at first full-term pregnancy (before and after 30 years old), breastfeeding (yes/no), menopausal status (yes/no), age at menopause (cut points: 47 and 54 years old among postmenopausal women only), mammographic exam during the previous follow-up period (yes/no), use of menopausal hormone therapy (ever, never and premenopausal women), previous family history of breast cancer (yes/no), previous personal history of benign breast disease (ever/never)

