UNIVERSITÉ PARIS-SUD 11
ÉCOLE DOCTORALE 420
Laboratoire : CESP - Épidémiologie environnementale des cancers

Année 2014

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS-SUD 11

Spécialité : Santé Publique - Épidémiologie

Présentée et soutenue publiquement

le 20 juin 2014

par

Claire DEMOURY

VARIATIONS GÉOGRAPHIQUES DE L'INCIDENCE DES LEUCÉMIES DE L'ENFANT ET ASSOCIATION AVEC L'EXPOSITION AUX RADIATIONS IONISANTES D'ORIGINE NATURELLE

Directeur de thèse : Monsieur Denis HÉMON

JURY

Madame Chantal GUIHENNEUC-JOUYAUX	Présidente
Monsieur Marc COLONNA	Rapporteur
Monsieur Denis ZMIROU-NAVIER	Rapporteur
Madame Isabelle DELTOUR	Examinatrice
Madame Agnès LEFRANC	Examinatrice
Monsieur Denis HÉMON	Directeur de thèse

Thèse préparée dans le cadre du Réseau doctoral en santé publique animé par l'Ecole des Hautes Études en Santé Publique (EHESP)

Adresse du laboratoire :
Centre de recherche en Epidémiologie et Santé des Populations
Equipe 6, Epidémiologie environnementale des cancers
Hôpital Paul Brousse, Bât. Inserm 15/16
16 Avenue Paul Vaillant-Couturier
94807 Villejuif Cedex

I. Remerciements

Je remercie d'abord, très sincèrement, mon directeur de thèse Denis Hémon pour m'avoir si bien encadrée pendant ces années de thèse. Merci pour son exceptionnelle disponibilité et toute l'attention qu'il a témoignée à mon travail et ce, malgré ses nombreuses responsabilités. Sa passion pour la recherche, sa rigueur et ses qualités scientifiques m'ont guidée tout au long de ce travail. Je lui suis extrêmement reconnaissante, ainsi qu'à Jacqueline Clavel, de m'avoir donné l'opportunité de travailler sur un sujet qui m'a passionnée.

Je voudrais également remercier les membres de mon jury pour leur implication dans l'évaluation de ce projet. Merci au Professeur Chantal Guihenneuc-Jouyaux qui m'a fait l'honneur d'en être la présidente, à Monsieur Marc Colonna et au Professeur Denis Zmirou-Navier pour leurs commentaires constructifs ainsi qu'à Mesdames Isabelle Deltour et Agnès Lefranc qui ont accepté d'être examinatrices.

Je remercie aussi les membres de l'IRSN qui ont contribué à ce travail : Sophie Ancelet, Christophe Debayle, Jérôme Guillevic, Géraldine Ielsch, Olivier Laurent et plus particulièrement Dominique Laurier, pour ses explications toujours extrêmement claires, la pertinence de ses questions pendant les réunions de travail et pour le temps qu'il a réussi à accorder à la relecture de mon manuscrit.

Je souhaite aussi remercier les membres de la société Géocible avec qui j'ai eu plaisir à travailler et plus particulièrement Mathieu Carrère pour tout l'intérêt qu'il porte au programme GEOCAP.

Je remercie Stéphanie Goujon pour tout ce qu'elle m'a appris pendant mon stage de master 2 et qui m'a si souvent servi pendant ma thèse. Sa grande disponibilité, sa bonne humeur et sa patience ont été très précieuses. Un grand merci également pour le temps qu'elle a consacré à la relecture de ma thèse.

Je n'oublie pas Pascale Bernillon, ma maître de stage à l'InVS, sans qui je n'aurais jamais eu l'idée de me lancer dans la recherche. Ses conseils et ses encouragements, à l'époque, ont été décisifs.

Merci à Michèle Paquet, Valérie Salmon et Olivier Bergada qui travaillent au CRIS de l'IRSN et ont grandement facilité mes recherches bibliographiques et, sans qui, la partie bibliographie de ma thèse aurait été bien peu conséquente !

Je tiens également à remercier Audrey Bourgeois et Evelyne Przybilski pour leur gentillesse et leur organisation dans les moindres petites tâches administratives.

Un grand merci à Olivier, pour son aide de dernière minute en informatique (et pas seulement !).

Je n'oublie évidemment pas toute l'équipe 6 du CESP au sein de laquelle il fait bon travailler et en particulier Claire, Alicia, Chloé, Thomas, Sofiène pour nos nombreux points sur l'actualité sportive, Audrey pour ses conseils de jeune docteur, Solène, Marie, Francesca, Alice, Jennifer, Céline, Lucie, Nicolas, Emilie N, Fabienne et tous les autres, arrivés plus récemment au sein de l'équipe... Un merci encore plus particulier à mes anciennes et actuelles collègues de bureau Solène, Jennifer et Alicia ou même de TMB, n'est-ce pas Marie ! Ces moments passés à l'Inserm et en dehors sont vraiment de très bons souvenirs !

Je remercie mes parents pour leur soutien et leurs encouragements durant cette thèse… et depuis toujours. Je remercie mes trois grandes sœurs Sandrine, Estelle et Agnès pour leur présence, et parce que, de près ou d'un peu plus loin, elles m'ont toujours soutenue. Je tiens à vous tous plus que tout.

Je remercie spécialement Florent qui me supporte depuis (presque !) un an et dont la présence à mes côtés a été essentielle ces derniers temps.

Merci à toute ma famille, cousins et cousines trop nombreux pour que je les cite, à mes colocataires, à mes amis picards, lillois ou parisiens…, qui sont présents autour de moi et qui, indirectement, ont largement contribué à l'aboutissement de ma thèse. Une dédicace particulière à mes deux « stateuses » d'amies Manon et Pauline !

Merci, enfin, à tous ceux qui m'ont un jour posé la question « Mais le radon, qu'est-ce que c'est ? » et ont eu la patience et l'intérêt d'écouter la réponse… !

II. Sommaire

I. Remerciements 5
II. Sommaire 7
III. Résumé 11
IV. Abstract 13
V. Productions scientifiques issues du travail de thèse 14
VI. Recueils de données utilisés et effectués par la doctorante dans le cadre de ce travail de thèse 15
VII. Liste des Tableaux 18
VIII. Liste des Figures 19
IX. Liste des abréviations 20

Partie II Problématique scientifique générale et objectifs 23
I. Introduction 24
A. Epidémiologie descriptive des leucémies de l'enfant 24
B. Facteurs de risque établis 26
C. Facteurs de risque suspectés 27
C.1. Facteurs infectieux 27
C.2. Facteurs environnementaux 28
II. Objectifs de la thèse et approche générale 32

Partie III Matériel et méthodes 35
I. Les cas 36
II. Les données démographiques 36
A. Découpage administratif 36
A.1. Les communes 36
A.2. Les bassins de vie 36
A.3. Les zones d'emploi 38
B. Estimations des populations 39
B.1. Données INSEE 39
B.2. Estimations du nombre de cas attendus par commune 41
III. Témoins 43
IV. Géolocalisation 44
A. Géocodage des adresses 44
B. Précision du géocodage 45
C. Validation du géocodage 48
V. Méthodes statistiques générales 49
A. Etude d'incidence 49
B. Etude cas-témoins 50
VI. Discussion 50

Partie IV Hétérogénéité spatiale de l'incidence des leucémies de l'enfant 53
I. Introduction 54
II. Matériel 55
A. Les cas 55
B. Population 55
III. Méthodes 56
A. Méthodes de détection d'une hétérogénéité globale (clustering) 56
A.1. Test de Potthoff et Whittinghill 56

A.2. Test de Rogerson ... 56
B. Méthodes de détection de clusters ... 57
B.1. Puissance des méthodes ... 58
B.2. Méthodes retenues ... 60
IV. Résultats ... 65
V. Discussion ... 70

Partie V Exposition aux radiations ionisantes d'origine naturelle ... 73
I. Introduction ... 74
A. La radioactivité ... 74
B. Les sources de radioactivité naturelle ... 74
B.1. Le radon ... 75
B.2. Les rayonnements telluriques ... 75
B.3. Les rayonnements cosmiques ... 76
B.4. Les eaux minérales et les aliments ... 76
C. Les doses et leur unité de mesure ... 76
D. Exposition aux radiations d'origine naturelle en France ... 78
II. Matériel ... 79
A. Données d'exposition au radon ... 79
A.1. Campagne nationale de mesures d'activité volumique (AV) ... 79
A.2. Potentiel d'exhalaison du radon (PR) ... 82
B. Données d'exposition aux rayonnements gamma telluriques et cosmiques (RTC) ... 84
B.1. Campagne nationale de mesures ... 84
B.2. Réseau de dosimétrie environnementale (dosimètres thermoluminescents TLD) ... 85
III. Méthodes ... 86
A. Estimation de l'exposition au radon ... 86
A.1. Score d'exposition ... 87
A.2. Cokrigeage ... 89
B. Estimation de l'exposition aux RTC ... 91
IV. Résultats ... 93
A. Exposition au radon ... 93
A.1. Relation entre activité volumique, caractéristiques de l'habitat et potentiel radon ... 94
A.2. Score d'exposition ... 97
A.3. Cokrigeage ... 100
B. Exposition aux RTC ... 102
C. Exposition au radon et aux RTC ... 103
V. Discussion ... 104

Partie VI Radiations ionisantes d'origine naturelle et leucémies de l'enfant ... 109
I. Introduction ... 110
A. Cohorte des survivants japonais des bombardements atomiques d'Hiroshima et Nagasaki ... 111
B. Les expositions médicales ... 112
B.1. Exposition à des fins thérapeutiques ... 113
B.2. Exposition à des fins diagnostiques ... 114
C. Les expositions professionnelles ... 115
C.1. Expositions professionnelles des parents et leucémies chez les enfants ... 115
C.2. Expositions professionnelles et leucémies de l'adulte ... 116
D. Les expositions environnementales ... 119
D.1. Expositions accidentelles ... 119
D.2. Proximité des sites nucléaires ... 120
D.3. Expositions aux radiations ionisantes d'origine naturelle ... 121
II. Matériel et méthodes ... 126
A. Sujets ... 127
A.1. Etude d'incidence ... 127

A.2. Etude cas-témoins 127
B. Exposition aux radiations 127
B.1. Exposition au diagnostic et exposition périnatale. 128
B.2. Exposition cumulée 130
B.3. Dose reçue à la moelle osseuse 130
C. Analyses statistiques 131
D. Prise en compte des facteurs de confusion 132
D.1. Indice de défaveur au niveau contextuel 132
D.2. Co-expositions 132
E. Puissance de l'étude 133
III. Résultats 133
A. Radon et leucémies de l'enfant 133
B. RTC et leucémies de l'enfant 140
C. Radon, RTC et leucémies de l'enfant 143
D. Puissance de l'étude 143
IV. Discussion 143
Partie VII Conclusion générale et perspectives 157
I. Conclusion générale 158
II. Perspectives 159
Partie VIII Bibliographie 163
Partie IX Annexes 179
Partie X Articles 181

III. Résumé

Les rayonnements ionisants sont un facteur de risque reconnu pour les leucémies chez l'homme pour des fortes doses d'exposition médicale ou accidentelle. Ce résultat a été démontré à travers plusieurs études, en particulier le suivi des survivants des bombardements atomiques d'Hiroshima et Nagasaki (Preston et al., 1994), mais également le suivi de populations traitées par radiothérapie (UNSCEAR, 2000) ou exposées in utero (Doll and Wakeford, 1997),(Wakeford and Little, 2003). Ces études concernent des expositions aiguës ou répétées au cours du temps à des niveaux de dose modérés voire forts. En revanche, l'hypothèse de l'existence d'un risque associé aux rayonnements ionisants à des niveaux d'exposition inférieurs, habituellement rencontrés dans l'environnement reste à démontrer. Notre travail propose d'évaluer l'hypothèse de l'existence d'une association entre les expositions environnementales aux radiations ionisantes d'origine naturelle et le risque de leucémie de l'enfant (LA) en utilisant des observations réalisées en France métropolitaine.

Les cas de leucémie inclus dans ce travail sont toutes les LA du Registre National des Hémopathies malignes de l'Enfant, qui enregistre l'ensemble des cas de moins de 15 ans diagnostiqués en France métropolitaine, sur la période étudiée.

Un premier travail a consisté à étudier la répartition spatiale de l'incidence des leucémies de l'enfant au niveau des 1 916 *bassins de vie* (BV) définis par l'INSEE. Des méthodes de détection de cluster ont été appliquées sur les 7 675 cas de leucémies de l'enfant diagnostiqués au cours de la période 1990-2006 afin d'identifier les zones potentiellement associées à un plus fort risque de leucémies aiguës de l'enfant. Cette étude, qui portait sur l'ensemble de la France métropolitaine n'a pas mis en évidence d'hétérogénéité spatiale des taux d'incidence des LA de l'enfant au cours de la période 1990-2006 au niveau des BV. Cependant, quelques clusters spatiaux ont été identifiés dans des lieux et périodes spécifiques. Bien que les niveaux de significativité de ces clusters ne soutiennent pas fortement l'existence de facteurs de risque localisés, les clusters peuvent montrer un léger impact de facteurs de risque partagés à l'échelle des BV, y compris dans le contexte d'expositions environnementales, qui nécessitent une enquête plus approfondie.

Pour tester l'hypothèse de l'existence d'une association entre l'exposition environnementale aux radiations ionisantes d'origine naturelle et l'incidence des leucémies de l'enfant, une étude d'incidence basée sur les 9 056 cas de LA de la période 1990-2009 a été réalisée. Cette étude a été complétée par une étude cas-témoins en population fondée sur les 2 763 cas de LA enregistrés sur la période 2002-2007 et un ensemble témoin de 30 000 sujets, construit en lien avec l'INSEE, constituant un échantillon contemporain représentatif de la population pédiatrique française. Dans cette approche, la géolocalisation des adresses des cas et des témoins ainsi que celle des sources d'exposition et leur caractérisation permet de définir les critères de l'intensité d'exposition aux facteurs d'intérêt et de les mettre en relation avec le statut cas vs témoins des sujets.

Les données de base concernant l'exposition à la radioactivité d'origine naturelle ont été produites par l'IRSN (Institut de Radioprotection et de Sûreté Nucléaire). Une cartographie du potentiel d'exhalation du radon émis par le sol ainsi qu'un échantillon national de 10 843 points de mesures localisés dans des habitations (dosimètres passifs mis en place durant 6 mois) ont permis d'estimer l'exposition résidentielle au radon au niveau de la commune et du domicile. L'exposition aux rayonnements gamma telluriques et cosmiques a été estimée par *zone d'emploi* à partir des mesures issues de la campagne nationale IRSN et de mesures environnementales réalisées par des dosimètres thermoluminescents dans approximativement 1 000 sites couvrant la France entière, dans un but de surveillance de la radioactivité ambiante. L'ensemble représente plus de 28 000 mesures réalisées à l'aide de dosimètres entre 1977 et 2012 et réparties sur le territoire français.

Notre étude n'a pas montré d'association entre les leucémies de l'enfant et les expositions au radon et aux rayonnements gamma d'origine tellurique et cosmique estimés au diagnostic et de façon cumulée pendant l'enfance. Elle avait une bonne puissance pour mettre en évidence les risques attendus d'après les modèles de risque actuels (UNSCEAR) issus des études sur les risques observés à forte dose. Ce travail n'étaye pas l'hypothèse qu'il existerait une association, directement observable sur le plan épidémiologique, entre les expositions aux rayonnements ionisants d'origine naturelle. Cette question reste néanmoins suffisamment importante et peu explorée pour mériter des études complémentaires de puissance analogue dans les pays où elle a été réalisée à grande échelle (Danemark, Grande-Bretagne, Suisse et France) et de réplication dans d'autres pays.

IV. Abstract

Ionizing radiation due to medical or accidental exposure to high doses is an established risk factor for leukemia in humans. This has been demonstrated through several studies; in particular the monitoring of survivors of the atomic bombings of Hiroshima and Nagasaki (Preston et al., 1994), but also of populations exposed therapeutically to radiations (UNSCEAR, 2000) or exposed in utero (Doll and Wakeford, 1997; Wakeford and Little, 2003). These studies involve acute or chronic exposures to moderate or high doses. However, the evidence remains to be demonstrated of a risk associated with exposure to ionizing radiation at lower levels, the ones usually encountered in the environment. Our work aims to evaluate the hypothesis of the existence of an association between natural background ionizing radiation and the risk of childhood leukemia (CL) using observations made in France.

Leukemia cases included in this study are all the CL recorded over the studied period in the National Registry of Childhood Haematological Malignancies, an exhaustive repository of all cases of patients younger than 15 years old in France.

First step was the study of the spatial distribution of the incidence of CL at the level of the 1,916 *Living Zone* (LZ) defined by INSEE. Cluster detection methods have been used on 7,675 cases of CL diagnosed during the period 1990-2006 to identify areas potentially associated with a higher risk of acute childhood leukemia. The study, covering the whole French territory, did not show any spatial heterogeneity of incidence of CL during the period at LZ level. However, some spatial clusters were highlighted in specific places and times. Although the levels of significance of these clusters do not strongly support the existence of risk factors, localized clusters can show a slight impact of risk factors shared across LZ, including contextual environmental exposures, that requires further investigation.

To test the hypothesis of the existence of an association between environmental exposure to ionizing radiation of natural origin and incidence of childhood leukemia, an incidence study based on 9,056 cases of CL for the period 1990-2009 was conducted. This study was complemented by a record-based cases-controls study based on the 2,763 cases of CL recorded over the 2002-2007 period and a control set of 30,000 subjects, gathered in collaboration with INSEE, constituting a representative sample of the contemporary French pediatric population. In this approach, localizations of cases and controls and exposure identifications were geocoded and compared to the status cases vs control population.

Data of exposure to natural background radiation were produced by the IRSN (Institute for Radiological Protection and Nuclear Safety). Mapping of the “potential radon exhalation emitted by the ground” and a national sampling of 10,843 measurement points located in dwellings (passive dosimeters in place for 6 months) were used to estimate residential exposure to radon at a level of granularity of cities and houses. Exposure to terrestrial gamma and cosmic rays was estimated by *zone d’emploi* based on a set of more than 28,000 measurements: environmental measurements made by thermoluminescent dosimeters in approximately 1,000 sites covering whole France in the context of environmental radioactivity monitoring and the data of the IRSN national campaign.

Our study did not show any association of childhood leukemia with exposures to radon and gamma radiation estimated nor at diagnosis nor cumulatively during childhood. However it had a good power to highlight the risks expected from current models of risk (UNSCEAR) built from studies on the observed high doses risks. If this work does not support the hypothesis that there is an association between exposure to ionizing radiation from natural sources observed and the incidence of childhood leukemia which may be directly observable at the epidemiologic level, this question remains important enough and not investigated enough to merit further complementary studies of analogous power in countries where it has been investigated so far (United-Kingdom, Denmark, Switzerland and France) and replication studies in countries where it has not been investigated.

V. Productions scientifiques issues du travail de thèse

Travaux relatifs à la thèse

Articles publiés

Goujon-Bellec S, **Demoury C**, Guyot-Goubin A, Hémon D, Clavel J. Detection of clusters of a rare disease over a large territory: performance of cluster detection methods. Int J Health Geogr. 2011 Oct 4;10:53.

Demoury C, Goujon-Bellec S, Guyot-Goubin A, Hémon D, Clavel J. Spatial variations of childhood acute leukaemia in France, 1990-2006: global spatial heterogeneity and cluster detection at 'living-zone' level. Eur J Cancer Prev.2012 Jul;21(4):367-74.

Laurent O, Ancelet S, Richardson DB, Hémon D, Ielsch G, **Demoury C**, Clavel J, Laurier D. Potential impacts of radon, terrestrial gamma and cosmic rays on childhood leukemia in France: a quantitative risk assessment. Radiat Environ Biophys. 2013 May;52(2):195-209

Demoury C, Ielsch G, Hémon D, Laurent O, Laurier D, Clavel J, Guillevic J. A statistical evaluation of the influence of housing characteristics and geogenic radon potential on indoor radon concentrations in France. J Environ Radioact. 2013 Sep 17;126C:216-225

Article en preparation

Association between natural background radiation and childhood leukaemia: analysis of the 1990-2009 incidence of childhood leukemia in France and of the 2002-2007 record-based case-control study GEOCAP.

Communications orales

Demoury C, Ielsch G, Laurent O, Laurier D, Hémon D, Clavel J. Role of residential radon in childhood leukaemia incidence: the GEOCAP program. Congress of the International Radiation Protection Association, 17 juin 2012, Glasgow.

Demoury C, Hémon D, Clavel J. Exposure to natural background radiation and association with childhood leukemia incidence in France. Rencontres scientifiques du réseau doctoral de l'EHESP, 10 avril 2014, Paris.

Communications affichées

Demoury C, Goujon-Bellec S, Guyot-Goubin A, Hémon D, Clavel J. Spatial variations of childhood acute leukaemia in France, 1990-2006: global spatial heterogeneity and cluster detection at 'living-zone' level. Congrès du Groupe des REgistres de Langue Latine, 2 juin 2011, Caen.

Goujon-Bellec S, **Demoury C**, Guyot-Goubin A, Hémon D, Clavel J. Detection of clusters of a rare disease over a large territory: performance of cluster detection methods. Congrès du Groupe des REgistres de Langue Latine, 2 juin 2011, Caen.

Travaux parallèles à la thèse

Articles publiés

Sermage-Faure C, **Demoury C**, Rudant J, Goujon-Bellec S, Guyot-Goubin A, Deschamps F, Hémon D, Clavel J. Childhood leukaemia close to high-voltage power lines--the GEOCAP study, 2002-2007. Br J Cancer. 2013 May 14;108(9):1899-906

VI. Recueils de données utilisés et effectués par la doctorante dans le cadre de ce travail de thèse

RNHE : Le présent travail de thèse a nécessité l'utilisation des données concernant les 9 056 cas de leucémies aiguës enregistrés dans le Registre National des Hémopathies malignes de l'Enfant (RNHE) sur la période 1990-2009 avec lequel nous avons étroitement collaboré tout au long de nos travaux, notamment pour prendre en compte les mises à jour successives du RNHE au fur et à mesure des recueils et validation d'années d'enregistrement.

INSEE : Pour estimer les populations par commune, année, âge et sexe nous avons utilisé les données disponibles sur le site de l'Institut National de la Statistique et des Etudes Economiques (INSEE) au niveau communal pour les années de recensement et départemental pour les autres.

IRSN : Nous avons travaillé en étroite collaboration avec l'Institut de Radioprotection et de Sûreté Nucléaire (IRSN) pour estimer l'exposition aux radiations ionisantes d'origine naturelle. Un groupe de travail a été mis en place avec des membres du Laboratoire d'Epidémiologie de l'IRSN (IRSN, PRP-HOM, SRBE, LEPID) et du Bureau d'étude et d'expertise sur la radioactivité naturelle (IRSN, PRP-DGE, SEDRAN, BRN). Pour le recueil des données sur les radiations gamma telluriques et cosmiques, nous avons aussi travaillé avec le Laboratoire de surveillance atmosphérique et d'alerte (IRSN, PRP-ENV, SESURE, LS2A). Le travail d'estimation de l'exposition aux radiations d'origine naturelle a porté sur le traitement et l'homogénéisation des différentes bases de données utilisées, la tabulation des résultats, les représentations graphiques afin de détecter d'éventuels éléments aberrants et les discuter avec les spécialistes concernés, la géolocalisation des mesures lorsqu'elles étaient accompagnées d'une adresse (cf géolocalisation automatique des adresses).

Dans ce groupe de travail, nous avons également travaillé sur une évaluation de risque des leucémies de l'enfant (LA) qui a fait l'objet d'une publication en 2013 (Laurent et al., 2013).

Lignes à haute-tension : Un autre volet du programme GEOCAP (GEOlocalisation des CAncers Pédiatriques, programme de recherche sur les cancers de l'enfant) auquel j'ai contribué pendant les deux premières années de cette thèse est l'exposition des sujets aux champs magnétiques dus à la proximité des lignes à haute tension. Un groupe de travail a été mis en place avec plusieurs experts de Rte (Réseau de transport d'électricité) afin d'évaluer au mieux l'exposition des sujets. Cette collaboration, à laquelle j'ai pris part, a été décrite dans une convention détaillant et précisant les tâches et responsabilités de chacune des deux entités dans nos travaux communs. Dans un premier temps le groupe de travail « exposition » de Rte a défini les méthodes de caractérisation des expositions liées à une adresse en général. Ce travail a été fait en concertation avec nous mais de façon totalement indépendante de toute analyse d'association exposition-leucémie de l'enfant. Dans un second temps, le groupe de travail de l'Inserm a réalisé les études d'association entre lignes à haute tension et LA basées sur la distance entre l'habitation d'un sujet et la ligne. La valeur du champ magnétique décroît rapidement avec la distance à la ligne. Il a donc été nécessaire de préciser la géolocalisation des sujets potentiellement exposés. Afin de déterminer très précisément l'exposition des sujets aux champs magnétiques induits par les lignes à haute tension, nous avons cherché à déterminer la localisation exacte du domicile (maison ou bâtiment) des sujets à partir d'examens photographiques des lieux (géolocalisation « au domicile »). Le géocodage « au domicile » a été réalisé en collaboration avec la société Géocible (Société d'études et de conseil spécialisée dans le géomarketing), à partir de bases de données (Google Map, Street View) permettant de rapprocher informations photographiques et adresses.

Ce travail a fait l'objet d'une première publication (Sermage-Faure, Demoury et al., 2013).

Géolocalisation automatique des adresses : La géolocalisation des adresses des sujets de l'étude GEOCAP et des mesures de la campagne nationale de l'IRSN a été faite en étroite collaboration avec l'entreprise Géocible. Nous nous sommes également impliqués de façon importante dans la conduite de cette collaboration à toutes les étapes. En effet, suite à

la détection d'une surévaluation de la précision de la géolocalisation des sujets GEOCAP, une priorité a été mise à l'amélioration du géocodage. Nous avons cherché à éliminer les erreurs et à déterminer une borne supérieure plus fiable des imprécisions de géocodage. En collaboration avec la société Géocible, nous avons élaboré et mis en place un système de détection d'erreur de géocodage afin d'éviter celles-ci en les repérant et les corrigeant. Pour cela, différents critères de fiabilité ont été mis en place (distance à la mairie, point appartenant au polygone de la commune, etc.). S'ils n'étaient pas vérifiés, le géocodage faisait alors l'objet d'une vérification et d'une correction si nécessaire. En parallèle, toutes les adresses ont été retraitées avec la base de donnée NAVTEQ mise à jour et incluant davantage d'informations sur la localisation des numéros le long des voies ainsi que la localisation de voies à l'intérieur de lieux-dits.

VII. Liste des Tableaux

Tableau 1. Type de géocodage obtenu pour les sujets GEOCAP et les mesures de radiations IRSN 47

Tableau 2. Distribution de la distance en mètres entre géocodage et géocodage « au domicile » 48

Tableau 3. Distribution statistique des bassins de vie français, en nombre de communes, taille de population et effectifs attendus de cas de leucémies de l'enfant....55

Tableau 4. Test de l'existence d'une sur-dispersion extra-poissonienne des taux d'incidence des leucémies de l'enfant en France métropolitaine (1990-2006) observés dans les 1 895 bassins de vie métropolitains....65

Tableau 5. Clusters les plus probables détectés par les méthodes SaTScan et FleXScan spatiales sur la période 1990-2006....66

Tableau 6. Clusters les plus probables détectés par la méthode SaTScan spatio-temporelle sur la période 1990-2006....66

Tableau 7. Clusters les plus probables détectés par les méthodes SaTScan et FleXScan pour les sous-périodes 1990-1994 ; 1995-2000 ; 2001-2006....69

Tableau 8. Pourcentages de variabilité intra et inter unités géographiques 92

Tableau 9. Distribution des mesures d'AV par classe de potentiel radon....94

Tableau 10. Activité volumique de radon et caractéristiques de l'habitat 96

Tableau 11. Distribution de l'estimation de l'exposition au radon 101

Tableau 12. Caractéristiques de l'exposition aux RTC à l'échelle des ZE....103

Tableau 13. Exposition aux RTC par classe d'exposition au radon....104

Tableau 14. Mobilité résidentielle et exposition au radon des sujets de l'étude cas-témoins Escale 129

Tableau 15. Association entre l'exposition au radon et l'incidence de toutes les leucémies aiguës 136

Tableau 16. Association entre l'exposition au radon estimée par cokrigeage et l'incidence des LAM par classe d'âge 137

Tableau 17. Association entre l'exposition au radon au diagnostic estimée par cokrigeage au domicile et l'incidence de toutes les leucémies aiguës 138

Tableau 18. Association entre l'exposition cumulée au radon[a] et l'incidence de toutes les leucémies aiguës....139

Tableau 19. Association entre l'exposition aux RTC et l'incidence de toutes les leucémies aiguës 141

Tableau 20. Association entre la dose à la moelle et l'incidence de toutes les leucémies aiguës 142

Tableau 21. Principales études rapportant un taux d'incidence chez les enfants exposés aux radiations (UNSCEAR 2006) 180

VIII. Liste des Figures

Figure 1. Taux d'incidence des leucémies aiguës en France (RNHE 2000-2004).....24

Figure 2. Taux de survie à 1 et 5 ans des enfants atteints de leucémies, selon différentes caractéristiques 25

Figure 3. Représentation des 1 916 bassins de vie français métropolitains 37

Figure 4. Géolocalisation d'une adresse dans différents cas de figure 46

Figure 5. Hypothèses de répartition non-uniforme des taux d'incidence des LA considérées dans cette étude.....59

Figure 6. FleXScan : zones de clusters potentiels 64

Figure 7. Clusters spatiaux détectés par les méthodes SaTScan et FleXScan 67

Figure 8. Pouvoir de pénétration des rayonnements 77

Figure 9. Distribution du logarithme de base 10 des mesures d'activité volumique (en Bq/m^3) 82

Figure 10. Cartographie du potentiel d'exhalaison du radon à l'échelle 1:1 000 000.....84

Figure 11. Semi-variogramme / Semi-variogramme expérimental 90

Figure 12. Relation entre les moyennes des mesures intérieures et extérieures dans les 167 ZE 93

Figure 13. Semi-variogramme empirique des 10 843 mesures d'AV 98

Figure 14. Moyennes d'AV par classes de score de potentiel radon.....99

Figure 15. Pourcentage de mesures d'activité volumique par classes de score de potentiel radon.....100

Figure 16. Distribution conjointe du logarithme de base 10 des estimations des AV à la mairie et à l'adresse des domiciles des témoins GEOCAP 101

Figure 17. Distribution conjointe du logarithme de base 10 des estimations des AV dans les ZE et dans les communes de résidence des témoins GEOCAP 102

Figure 18. Cartographie de l'exposition aux RTC à l'échelle des zones d'emploi 103

IX. Liste des abréviations

ASN : Autorité de Sûreté Nucléaire.

AV : Mesure d'activité volumique.

BEIR : Committee to Assess Health Risks from Exposure to Low-Levels of Ionising Radiation.

Bq : Becquerel (Bq/m^3 : nombre de désintégrations par m^3 et par seconde).

BV : Bassin de Vie, unité géographique regroupant des communes et défini par l'INSEE comme le plus petit territoire sur lequel les habitants ont accès aux équipements de la vie courante (supermarché, banque, cinéma,…) et à l'emploi.

CEM-EBF : Champs électriques et magnétiques à extrêmement basse fréquence.

CIRC : Centre International de Recherche sur le Cancer / **IARC :** International Agency for Research on Cancer.

CNPE : Centrale Nucléaire de Production d'Electricité.

CPP : Cluster le plus probable.

DDASS : Direction Départementale des Affaires Sanitaires et Sociales.

DGS : Direction Générale de la Santé.

EAR : Excess Absolute Risk (risque chez les sujets exposés – risque chez les sujets non exposés).

ERR : RR – 1 où RR est le risque relatif associé à une exposition.

GEOCAP : GEOlocalisation des CAncers Pédiatriques, programme de recherche sur les cancers de l'enfant dirigé par Jacqueline Clavel (Equipe 6 du CESP – U1018 Inserm).

Géocible : Société d'études et de conseil spécialisée dans le géomarketing.

IC95% : intervalle de confiance à 95%.

IGN : Institut Géographique National.

INSEE : Institut National de la Statistique et des Etudes Economiques.

InVS : Institut de veille sanitaire.

IPSN : Institut de Protection et de Sûreté Nucléaire (devenu l'IRSN).

IRSN : Institut de Radioprotection et de Sûreté Nucléaire.

LA : Leucémie aiguë de l'enfant.

LAL : Leucémie aiguë lymphoblastique de l'enfant.

LAM : Leucémie aiguë myéloblastique de l'enfant.

LSS : Life Span Study (cohorte des survivants des bombardements d'Hiroshima et Nagasaki).

O (resp. E) : Nombre de cas observés (resp. attendus).

OR : OR = [xx – xx] donne l'odds-ratio d'une association entre deux caractéristiques dichotomiques et son intervalle de confiance au risque 5%.

PR : Potentiel d'exhalaison du radon.

RNHE : Registre National des Hémopathies malignes de l'Enfant.

RNTSE : Registre National des Tumeurs Solides de l'Enfant.

RR : Risque Relatif.

RTC : Rayonnements gamma telluriques et cosmiques.

Rte : Réseau de transport d'électricité.

SIR : Ratio d'incidence standardisé.

Sv : Sievert.

TLD : Dosimètres thermoluminescents. Dans le document, cela fait référence aux mesures issues du réseau de dosimétrie environnementale du Laboratoire de surveillance atmosphérique et d'alerte (LS2A) de l'IRSN.

UNSCEAR : United Nations Scientific Committee on the Effects of Atomic Radiation.

ZE : Zone d'Emploi, unité géographique regroupant des communes et définie comme un espace géographique à l'intérieur duquel la plupart des actifs résident et travaillent, et dans lequel les établissements peuvent trouver l'essentiel de la main d'œuvre nécessaire pour occuper les emplois offerts.

Partie II Problématique scientifique générale et objectifs

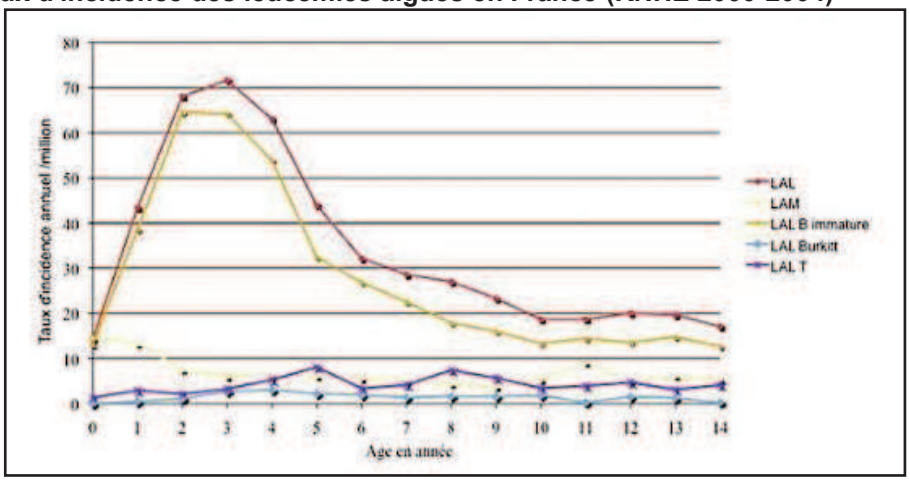
I. Introduction

A. Epidémiologie descriptive des leucémies de l'enfant

Les informations d'épidémiologie descriptive disponibles sur les leucémies de l'enfant en France métropolitaine sont issues de leur enregistrement depuis 1990 par le RNHE (Registre National des Hémopathies Malignes de l'Enfant).

Les hémopathies malignes représentent 41 % de la totalité des cancers avant l'âge de 15 ans en France (Lacour et al., 2010). Parmi elles, les leucémies aiguës (LA) sont majoritaires (29%) avec environ 470 cas par an. Il s'agit de maladies relativement rares avec un taux d'incidence brut de 41,7 cas par million et par an. Deux sous-types de leucémies se distinguent selon l'aspect cytologique lié à la lignée de cellules à partir de laquelle elles émergent : les leucémies aiguës lymphoblastiques (LAL) et les leucémies aiguës myéloblastiques (LAM). Les LAL représentent environ 80 % des cas de leucémie aiguë, et plus de 80 % d'entre elles sont des LAL B dites communes. La distribution des LAL B décrit un pic maximum vers l'âge de 2-3 ans, plus marqué chez les garçons que chez les filles, et encore nettement visible un peu après 6 ans. Le pic ne concerne que les LAL B communes. Les LAM représentent moins de 20% des cas de LA avec un pic d'incidence avant l'âge de un an (Figure 1).

Figure 1. Taux d'incidence des leucémies aiguës en France (RNHE 2000-2004)



Source : site web du RNHE : http://www.u754.idf.inserm.fr/page.asp?page=4611

Les pourcentages de survie observés en France métropolitaine sont proches de ceux obtenus dans les pays occidentaux. Pour la période 1990-2000, des différences de survie selon l'âge au diagnostic sont observées (Figure 2). Le taux de survie à 5 ans est actuellement proche de 90% pour les LAL, et de 65% pour les LAM. La trisomie 21 est un facteur pronostic défavorable dans les cas de LAL.

Figure 2. Taux de survie à 1 et 5 ans des enfants atteints de leucémies, selon différentes caractéristiques
(cas de leucémies diagnostiqués en France métropolitaine sur la période 1990-2000)

	1 an		5 ans	
	%	p	%	p
LAL (N=3960)				
Sexe				
Garçons	94 [93-95]		81 [79-82]	0.07
Filles	95 [94-96]		83 [81-85]	
Age au diagnostic				
< 1 an	77 [69-84]	$<10^{-4}$	48 [39-57]	$<10^{-4}$
1-4 ans	96 [95-97]		87 [85-88]	
5-9 ans	95 [94-96]		83 [81-85]	
10-14 ans	91 [89-94]		72 [69-76]	
Trisomie 21				
Non	94 [94-95]		82 [81-83]	$<10^{-2}$
Oui	92 [86-99]		66 [54-78]	
LAM (N=809)				
Sexe				
Garçons	78 [74-82]		56 [51-61]	
Filles	77 [73-81]		59 [55-64]	
Age au diagnostic				
< 1 an	61 [53-70]	$<10^{-4}$	45 [37-54]	$<10^{-3}$
1-4 ans	80 [75-84]		62 [56-68]	
5-9 ans	84 [78-89]		61 [54-68]	
10-14 ans	81 [75-86]		57 [51-64]	
Trisomie 21				
Non	79 [76-81]	0.06	58 [54-61]	
Oui	67 [55-80]		56 [42-69]	

Source : site web du RNHE : http://www.u754.idf.inserm.fr/page.asp?page=4611

Variations temporelles de l'incidence

Une analyse récente sur la période 1990-2007 a suggéré une augmentation de l'incidence des LAL (+0,48% IC95% [0,02% ; 0,95%]), notamment chez les filles les plus âgées (Goujon-Bellec et al., 2013). Cette augmentation qui semblait limitée à la période 1990-2001 pourrait être due à l'amélioration des méthodes de recueil des données et d'enregistrement de ces hémopathies et/ou à d'autres causes qui restent à identifier.

Variations spatiales de l'incidence

En France, une hétérogénéité géographique de l'incidence des LA sur la période 1990-1994 modérée et à la limite de la significativité a été observée à l'échelle des communes (surdispersion de 1,2%, p=0,06), tandis qu'aucune hétérogénéité n'était observée sur la période 1995-2000 (Bellec et al., 2006). A l'échelle des départements, une hétérogénéité modérée a été observée sur la période 1990-2004 (surdispersion de 18,6%, p=0,10) (Faure et al., 2009).

B. Facteurs de risque établis

Certains facteurs iatrogènes (chimiothérapie et radiothérapie) et certaines anomalies chromosomiques constitutionnelles sont clairement associés à un risque accru de leucémie. Ainsi, avant l'âge de 5 ans, les enfants porteurs d'une trisomie 21 ont un risque environ 50 fois supérieur de développer une leucémie par rapport aux enfants de la population générale. Entre 2 et 3% des enfants atteints de leucémie sont porteurs d'une trisomie 21 (Stiller, 2004). Certaines maladies génétiques rares entraînent également une augmentation du risque : l'ataxie-télangiectasie, la maladie de Fanconi (LAM), l'anémie de Diamond-Blackfan (LAM) (Stiller, 2004).

Des études d'association génome entier montrent des associations avec des polymorphismes de 2 gènes, *ARID5B* et *IKZF1,* impliqués dans la différentiation des lymphocytes B (Orsi et al., 2012; Papaemmanuil et al., 2009).

Le rôle des radiations ionisantes à forte dose est établi pour la plupart des cancers de l'enfant (UNSCEAR, 2000). L'exposition externe aux rayonnements ionisants est classée « cancérigène certain » par le Centre International de Recherche sur le Cancer (CIRC) (IARC, 2000).

C. Facteurs de risque suspectés

C.1. Facteurs infectieux

L'origine virale des leucémies de l'enfant est suspectée depuis de nombreuses années. L'hypothèse d'une infection pouvant être associée à ce processus oncogène chez l'homme a été avancée compte tenu de l'origine virale des leucémies chez certaines espèces animales, de l'implication du virus d'Epstein-Barr dans la genèse du lymphome de Burkitt chez l'enfant (Peh et al., 2004) et de l'existence d'un certain nombre de clusters espace-temps très marqués de cas de leucémies de l'enfant (Kinlen, 1988).

Des hypothèses impliquant un agent infectieux survenant soit *in utero* (Smith, 1997; Smith et al., 1998) soit dans la petite enfance (Kinlen, 1988; Kinlen, 1995) ont été proposées. Selon Kinlen (Kinlen, 1988), une population faiblement exposée à un agent infectieux banal du fait d'une résidence dans une zone rurale isolée, pourrait, secondairement à un important brassage de population, se trouver en contact avec cet agent infectieux. De jeunes enfants, issus de cette population infectée, pourraient alors développer une réponse immunitaire anormale sous forme de leucémies. L'incidence des leucémies est apparue significativement plus élevée dans les communes isolées ayant connu les mouvements de population les plus importants. Une étude réalisée au sein de l'équipe[1] a mis en évidence une association entre l'incidence des leucémies de l'enfant et les mouvements de population du lieu de résidence à la naissance (Rudant et al., 2006) ou au diagnostic dans les communes isolées (Bellec et al., 2008). Cela est concordant avec les conclusions d'une récente méta-analyse de Kinlen qui n'a pour autant inclus qu'un nombre limité d'études (Kinlen, 2012). Greaves a fait l'hypothèse d'un mécanisme en deux temps (Greaves, 1988). Une première mutation, liée à une infection, interviendrait *in utero*. La seconde mutation, plus tardive dans la petite enfance et correspondant au pic d'incidence observé, serait en rapport avec une réponse exagérée à une infection banale, liée à l'immaturité du système immunitaire. Pour l'instant, aucun agent

[1] Centre de recherche en Épidémiologie et santé des Populations, Equipe 6 : Epidémiologie environnementale des Cancers – UMRS 1018

infectieux spécifique n'a été identifié. Les infections banales des premiers mois de la vie semblent jouer un rôle majeur dans la maturation du système immunitaire. En l'absence de stimulation suffisante à un âge précoce, les cellules lymphoïdes pré leucémiques qui apparaissent *in utero* chez environ 1% des fœtus encourraient un risque accru de transformation leucémique au cours d'épisodes infectieux de la petite enfance (Greaves, 2006). Plusieurs études montrent une association négative entre les leucémies de l'enfant et un nombre élevé d'infections banales avant un an ou la mise en collectivité précoce. Des études épidémiologiques ont porté sur l'effet de l'impact d'une immunisation précoce des enfants sur le risque de leucémies avec des approches très diverses. Certaines études ont étudié les marqueurs d'infections dans la petite enfance comme la mise en collectivité précoce (crèche) (Gilham et al., 2005; Perrillat et al., 2002), le nombre d'enfants dans la fratrie (Infante-Rivard et al., 2000; Jourdan-Da Silva et al., 2004), le brassage de population et les contacts interpersonnes des parents dans leur travail (Kinlen and Bramald, 2001) ou l'allaitement maternel (Kwan et al., 2004b). Le plus souvent, ces études ont montré une diminution du risque de leucémies en cas d'immunisation précoce mais d'autres ont obtenu des résultats contradictoires (Chan et al., 2002; Dockerty et al., 1999b; Petridou et al., 2001).

C.2. Facteurs environnementaux

C.2.a. Consommation parentale

L'association de l'incidence des leucémies de l'enfant avec les consommations de tabac, d'alcool et d'un certain type d'alimentation au cours de la grossesse a été suspectée. Il n'y a pas de forte évidence épidémiologique pour un lien avec le tabagisme de la mère au cours de la grossesse (Bonaventure et al., 2012; Brondum et al., 1999; Menegaux et al., 2005; Severson et al., 1993; Shu et al., 1996) mais des évidences concordantes sur une association des LA avec la consommation de tabac du père (Rudant et al., 2008). Le rôle de la consommation d'alcool au cours de la grossesse a donné lieu à des résultats divergents (Bonaventure et al., 2012; Infante-Rivard et al., 2002; Menegaux et al., 2005; Petridou et al., 1997; Severson et al., 1993; Shu et al., 1996) ainsi que les études sur l'alimentation au cours

de la grossesse (Ross et al., 1996) ou au cours de l'enfance (Kwan et al., 2004a; Peters et al., 1994). Une association avec la consommation régulière de café pendant la grossesse a été observée dans deux études de l'équipe (Bonaventure et al., 2013; Menegaux et al., 2005).

C.2.b. Expositions aux pesticides

Des associations positives ont été observées entre la leucémie infantile et l'exposition résidentielle aux pesticides pendant la grossesse ou l'enfance (Lowengart et al., 1987; Ma et al., 2002; Menegaux et al., 2006; Rudant et al., 2007b) et confirmées par trois récentes méta-analyses (Turner et al., 2010; Van Maele-Fabry et al., 2011; Vinson et al., 2011). Cependant des travaux complémentaires sont nécessaires pour confirmer ces résultats basés sur des questionnaires et évaluer plus en détail les spécificités et sous-groupes toxicologiques des pesticides (Zahm and Ward, 1998). Une méta-analyse conclut à une association positive avec l'exposition professionnelle maternelle et négative avec l'exposition paternelle (Wigle et al., 2009). Cela est retrouvé par Vinson et al.

C.2.c. Expositions aux hydrocarbures

Le benzène est leucémogène chez l'adulte (IARC, 1982). Les études sur le trafic automobile et le risque de LA, encore peu nombreuses, sont majoritairement en faveur d'un risque de leucémie accru à proximité immédiate des routes à fort trafic (Crosignani et al., 2004; Feychting et al., 1998; Heck et al., 2013; Nordlinder and Jarvholm, 1997; Savitz and Feingold, 1989; Visser et al., 2004). Le registre de cancers lombard rapporte un risque de leucémie significativement lié au niveau estimé de benzène produit par le trafic automobile au lieu de résidence au diagnostic avec un OR de 3,9 (IC95% [1,4 ; 11,3]) associé aux expositions au benzène estimées à plus de 10 µg/m^3 (Crosignani et al., 2004). Néanmoins, d'autres études allemande et italienne n'ont pas retrouvé cette association (Badaloni et al., 2013; Raaschou-Nielsen et al., 2001) comme une étude conduite en Californie avec la densité de trafic (Reynolds et al., 2004; Reynolds et al., 2002; Von Behren et al., 2008) ; ceci est confirmé par la revue de Raaschou-Nielsen et Reynolds (Raaschou-Nielsen and

Reynolds, 2006). Du point de vue des expositions au benzène, nous pouvons en rapprocher les résultats de deux études cas-témoins successives réalisées dans l'équipe, convergentes et évoquant l'augmentation possible du risque de leucémie à proximité immédiate de garages ou de stations-service (OR entre 2 et 4, concernant 3 à 6% de la population pédiatrique) (Brosselin et al., 2009; Steffen et al., 2004) ainsi que les résultats d'une autre étude rapportant une augmentation de risque à moins de 100 mètres des stations-service (Harrison et al., 1999). Dans une étude plus récente (Amigou et al., 2011), les LA étaient significativement associées à des estimations de la concentration de NO_2 lié au trafic sur le lieu de résidence supérieure à 27,7 µg/m^3 par rapport à une concentration inférieure à 21,9 µg/m^3 (OR = 1,2 IC95% [1,0 ; 1,5]) et avec la présence d'une route à fort trafic à moins de 500 mètres (OR = 2,0 IC95% [1,0 ; 3,6]). Les LA étaient également associées à une forte densité de routes à fort trafic dans les 500 mètres (OR = 2,2 IC95% [1,1 ; 4,2]) avec une tendance linéaire positive significative de l'association avec la longueur totale de la route à fort trafic à moins de 500 m.

C.2.d. Expositions aux radiations non ionisantes

Les expositions aux champs électriques et magnétiques à extrêmement basse fréquence (CEM-EBF) ont fait l'objet d'un nombre important de travaux. Les méta-analyses (Ahlbom et al., 2000; Greenland et al., 2000) ont amené le CIRC à classer l'exposition aux champs magnétiques de 0,4 µT ou plus comme « Cancérogène possible » (2B) (IARC, 2002). Cette exposition doublerait le risque de leucémie de l'enfant et concernerait environ 1% de la population pédiatrique. Dans l'équipe, dans une étude cas-témoins basée sur des données géolocalisées (GEOCAP) un OR de 1,7 (IC95% [0,9 ; 3,6]) a été trouvé pour les sujets résidant à moins de 50 m d'une ligne à très haute tension (225 kV ou 400 kV) (Sermage-Faure et al., 2013). 9 cas et 60 témoins étaient concernés.

Les expositions aux UV de l'enfance accroissent le risque de mélanome de l'enfant et de l'adulte. Leur implication dans d'autres risques de cancer a très peu été décrite. Le rôle de l'exposition aux UV dans le risque d'hémopathies malignes de l'enfant a été peu étudié.

Deux récentes études américaines ont suggéré un lien négatif avec les LAL de l'enfant (Lombardi et al., 2013; Musselman and Spector, 2011) tandis qu'une étude conduite en Finlande a rapporté un lien positif (Timonen et al., 2007).

Les données sur les radiofréquences sont encore extrêmement limitées. Une étude anglaise sur les antennes-relais fréquemment mises en cause par le public en santé publique amène plutôt à écarter la vraisemblance d'un lien avec les leucémies de l'enfant (Elliott et al., 2010).

C.2.e. Expositions aux radiations ionisantes

Le rôle des radiations ionisantes à forte dose est établi pour la plupart des cancers de l'enfant et notamment pour les leucémies. Concernant le rôle des radiations ionisantes à faible dose, les travaux les plus nombreux ont concerné le risque de leucémie à proximité des sites nucléaires et les radiations d'origine naturelle.

Dans la recherche des causes possibles de ces cancers, la plupart des études concernant les sites nucléaires, qu'elles aient été locales et postérieures à la découverte de clusters ou multi-sites et planifiées *a priori*, ont abouti à des résultats ne permettant ni d'affirmer l'existence d'une incidence plus élevée au voisinage des sites nucléaires, ni de pouvoir l'exclure de façon certaine (Laurier and Bard, 1999).

En France, 58% de l'exposition aux radiations ionisantes est d'origine naturelle (3 mSv en moyenne par an : 34% due au radon, 11% due aux rayons gamma d'origine tellurique, 7% due aux rayons gamma d'origine cosmique et 6% due à l'ingestion d'eau et de nourriture). La dose individuelle correspondante est comprise entre 2 et 3 mSv par an en France. Si l'on ajoute l'exposition aux sources artificielles de rayonnements ionisants, la dose atteint 4 mSv. On parle d'exposition à faible dose.

Des études épidémiologiques sur des groupes de personnes exposées à des doses modérées et fortes de rayonnements ionisants ont montré que la leucémie infantile était particulièrement sensible aux irradiations (BEIRVII, 2006a; ICRP, 2008; UNSCEAR, 2008). Cela a suscité beaucoup d'intérêt dans l'étude du risque de leucémie infantile résultant de

faibles doses de rayonnement reçues *in utero* ou dans les premières années de vie, par exemple, l'exposition médicale à des fins diagnostiques ou l'exposition aux radiations ionisantes d'origine naturelle (Wakeford, 2008). Récemment, l'attention accordée à ce sujet a augmenté, en partie à cause des rapports de taux d'incidence de la leucémie plus élevés chez les enfants vivant à proximité de certaines installations nucléaires (Laurier et al., 2008) mais aussi en raison de la fréquence croissante des méthodes modernes de radiographie médicale.

Les données de la littérature sur le lien entre radiations ionisantes et leucémie de l'enfant seront développées de façon plus précise en introduction de la Partie VI.

II. Objectifs de la thèse et approche générale

Les rayonnements ionisants sont un facteur de risque reconnu pour les leucémies chez l'homme pour des fortes doses d'exposition médicale ou accidentelle. Ce résultat a été démontré à travers plusieurs études, en particulier le suivi des survivants des bombardements atomiques d'Hiroshima et Nagasaki dans le cadre de la *Life Span Study* (LSS) (Preston et al., 1994), mais aussi celui de populations traitées par radiothérapie (UNSCEAR, 2000) ou exposées *in utero* (Doll and Wakeford, 1997; Wakeford and Little, 2003). Ces études concernent des expositions aiguës ou répétées au cours du temps à des niveaux de dose modérés à forts.

A partir des observations recueillies dans des populations ayant reçues de fortes doses le plus souvent dans un intervalle de temps bref (forts débits de dose) des modèles de risque « de référence », qui permettent des extrapolations à faibles doses très utiles pour établir des normes de radioprotection, ont été établis par l'UNSCEAR (United Nations Scientific Committee on the Effects of Atomic Radiation) et le BEIR (Committee to Assess Health Risks from Exposure to Low-Levels of Ionising Radiation).

La cancérogénicité établie des radiations ionisantes et ces modèles sous-tendent l'hypothèse de l'existence d'un risque de leucémies chez l'enfant lié aux rayonnements ionisants à des niveaux d'exposition inférieurs, habituellement rencontrés dans

l'environnement et délivrés de façon continue dans le temps. Cette hypothèse reste à éprouver par des observations épidémiologiques directes. S'inscrivant dans le contexte scientifique général décrit précédemment, notre projet de recherche propose une approche épidémiologique de l'hypothèse de l'existence d'une association entre les expositions environnementales aux radiations ionisantes d'origine naturelle et le risque de leucémie de l'enfant en France métropolitaine.

Dans un premier temps et sans *a priori* sur un facteur de risque, un travail introductif a porté sur l'étude des variations spatiales de l'incidence des leucémies en France, en particulier la recherche d'une hétérogénéité spatiale globale (« clustering ») et la recherche d'excès localisés d'incidence (« clusters »). La détection d'agrégats localisés est à l'origine du développement de nombreuses méthodes statistiques dont l'une, très répandue, développée par M.Kulldorff et N.Nagarwalla en 1995 (Kulldorff and Nagarwalla, 1995) et d'autres, plus récentes, comme celles développées par T. Tango et K. Takahashi en 2005 (Tango and Takahashi, 2005). Ces méthodes sont basées sur des fenêtres mobiles de forme circulaire, elliptique ou plus irrégulière et permettent de localiser des clusters. La présence d'une hétérogénéité spatiale peut refléter une propension générale de la maladie à survenir de façon agrégée dans l'espace, être secondaire à la présence locale d'expositions cancérogènes, ou encore témoigner des fluctuations d'incidence attendues sous l'hypothèse d'une distribution poissonienne des cas.

L'objectif principal de la thèse est de contribuer à répondre à la question de l'existence d'une association entre l'incidence des leucémies de l'enfant et l'exposition environnementale aux radiations ionisantes d'origine naturelle (radon et rayonnements gamma d'origine tellurique et cosmique) en examinant ce qu'il en est en France métropolitaine sur la période 1990-2009. Ce travail s'appuie principalement sur une étude d'incidence et, de façon secondaire, sur une étude cas-témoins pour la période 2002-2007. Elle repose sur des données de mesures effectuées par l'Institut de Radioprotection et de Sûreté Nucléaire (IRSN). Avant cette analyse, un premier travail a consisté à étudier la répartition spatiale de l'incidence des leucémies de l'enfant. Des méthodes de détection de

clusters ont été appliquées sur les données de leucémies de l'enfant au cours de la période 1990-2006 afin d'identifier les zones potentiellement associées à un plus fort risque de leucémies aiguës de l'enfant en France.

Partie III Matériel et méthodes

I. Les cas

Les données de leucémies aiguës proviennent du Registre National des Hémopathies Malignes de l'Enfant (RNHE), qui inclut en continu depuis 1995, et de manière rétrospective depuis 1990, tous les enfants présentant une hémopathie maligne ou à la limite de la malignité (leucémies, lymphomes et autres hémopathies), âgés de moins de 15 ans au diagnostic et résidant en France métropolitaine. Pour chaque cas, le registre fournit notamment la date de naissance, la date du diagnostic, le diagnostic et le code INSEE de la commune où le cas a été diagnostiqué (lieu de résidence de l'enfant au moment du diagnostic).

Sur la période 1990-2009, 9 056 cas de LA ont été diagnostiqués en France dont 7 434 cas de LAL et 1 465 cas de LAM.

II. Les données démographiques

A. Découpage administratif

A.1. Les communes

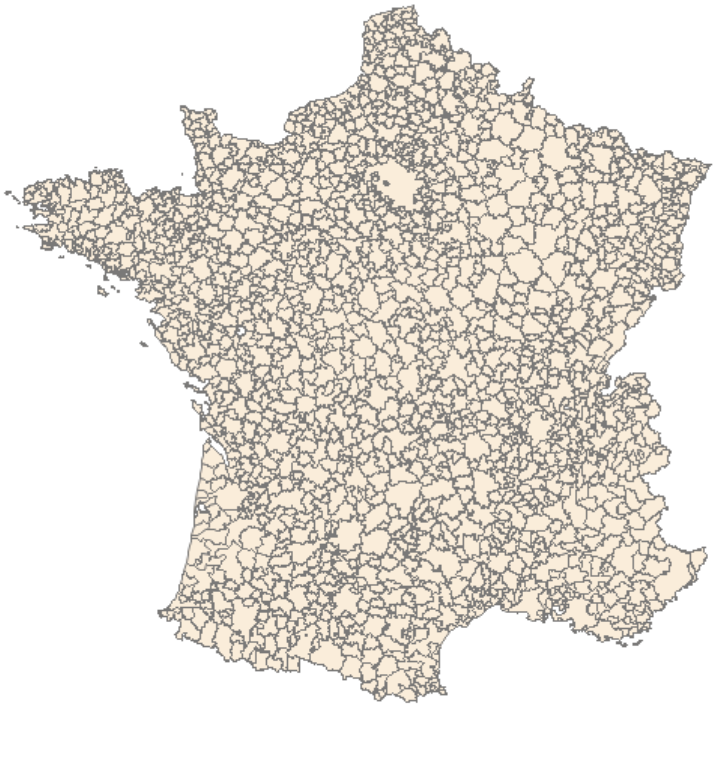
La commune est la plus petite subdivision administrative française. Au 1er janvier 2010, on comptait 36 682 communes, dont 36 570 en métropole.

A.2. Les bassins de vie

Les *bassins de vie* (BV) sont des unités géographiques créées en 2003 par l'INSEE pour faciliter la compréhension de la structuration du territoire de la France métropolitaine. Le *bassin de vie* est un ensemble de communes défini comme le plus petit territoire sur lequel les habitants ont accès aux équipements de la vie courante (supermarché, banque, cinéma,…) et à l'emploi. Le découpage de la France en BV est basé sur une maille initiale, les bassins de services intermédiaires, issue de la carte des « territoires vécus ». Cette carte, largement diffusée par l'INSEE et la Datar superpose le ZAUER (zonage en aires urbaines et aires d'emploi) et la manière dont les communes du pays sont reliées à des

pôles de services intermédiaires. Ces pôles correspondent à une commune disposant d'au moins neuf services parmi les supermarchés, drogueries, librairies, magasins de chaussures, magasins de meubles, magasins d'électroménager, magasins de vêtements, centres de secours, collèges, gendarmeries ou commissariats de police, ambulances, banques, cabinets de dentistes, kinésithérapeutes, vétérinaires, notaires. Une fois les pôles déterminés, les flux issus de l'Inventaire Communal ont servi à définir les bassins de services intermédiaires autour de chaque pôle. L'Inventaire Communal, un inventaire complet des équipements, commerces et services au niveau de chaque commune, réalisé en mars 1998 auprès de toutes les mairies de France, permet de connaître la densité d'implantation et la fréquentation des équipements. Jugés parfois de trop petite taille, les bassins de services intermédiaires ont été regroupés entre eux, via des calculs de scores, pour former les BV. La France est ainsi découpée en 1 916 *bassins de vie* (Figure 3), parmi lesquels une grande majorité (1 745) s'articule autour de localités de moins de 30 000 habitants.

Figure 3. Représentation des 1 916 bassins de vie français métropolitains



En 2012, suite à la rénovation d'autres zonages (unité urbaines, *zones d'emploi* 2010 et aires urbaines 2010), le découpage en BV a été rénové. Une différence notable concernant les 1 666 *bassins de vie* 2012 par rapport aux *bassins de vie* précédents est la disparition totale de toute notion d'emploi dans leur définition. Seule la notion d'accessibilité aux équipements est prise en compte.

Dans la présente étude, le découpage 2003 a été considéré.

A.3. Les zones d'emploi

Une *zone d'emploi* (ZE) est définie comme un espace géographique regroupant des communes à l'intérieur duquel la plupart des actifs résident et travaillent, et dans lequel les établissements peuvent trouver l'essentiel de la main d'œuvre nécessaire pour occuper les emplois offerts. Ce zonage est défini à la fois pour la France métropolitaine et les DOM. Le découpage actualisé se fonde sur les flux de déplacements domicile-travail des actifs observés lors du recensement de 2006. La liste des communes est celle donnée par le Code Officiel Géographique. La constitution des *zones d'emploi* repose sur l'agrégation par étape des communes selon l'intensité des déplacements domicile-travail.

À chaque étape :

- les flux des déplacements domicile-travail les plus importants entre unités (communes ou groupes de communes agrégées précédemment) et communes non encore agrégées, sont repérés ;

- si l'unité constituée dépasse un certain seuil en termes de nombre d'emplois, cette unité appelée alors pôle d'emploi peut agréger d'autres communes mais ne peut l'être avec une autre zone en constitution.

Le principe est donc de maximiser parmi l'ensemble des actifs en emploi d'une zone la proportion de « stables », c'est-à-dire les actifs travaillant dans leur zone de résidence.

La constitution des nouvelles *zones d'emploi* a respecté un certain nombre de contraintes relatives notamment à la taille des territoires (superficie, population, emplois) et à leur géographie (pas de fractionnement infra-communal, contiguïté des communes, partition

entière de la France, etc.), mais aussi à la mobilité des actifs (distance maximale domicile-travail). Pour mieux circonscrire les marchés locaux du travail, les *zones d'emploi* peuvent désormais dépasser les limites administratives régionales, même si celles-ci ont été le plus souvent respectées. La France métropolitaine est composée de 304 *zones d'emploi*.

Le découpage en *zones d'emploi* constitue une partition du territoire adaptée aux études locales sur le marché du travail. Le zonage définit aussi des territoires pertinents pour les diagnostics locaux et peut guider la délimitation de territoires pour la mise en œuvre des politiques territoriales initiées par les pouvoirs publics ou les acteurs locaux.

B. Estimations des populations

B.1. Données INSEE

Age

L'âge est la durée écoulée depuis la naissance. Il peut être calculé selon deux définitions :

- l'âge en années révolues est l'âge atteint au dernier anniversaire.

- l'âge atteint dans l'année (ou âge par génération ou encore âge en différence de millésimes) est la différence entre l'année de l'événement et l'année de naissance et correspond à l'âge en années révolues au 31 décembre de l'année considérée.

Par exemple, à la date du 1er janvier 2006, un individu né le 10 octobre 1925 est âgé de 81 ans en âge atteint dans l'année et de 80 ans si l'on considère son âge en années révolues.

B.1.a. Les données des recensements

Les populations des communes ont été dénombrées lors des recensements nationaux des 5 mars 1990 et 8 mars 1999 avec un comptage exhaustif. Les résultats de ces recensements sont présentés en utilisant l'âge atteint au cours de l'année du recensement.

Depuis janvier 2004, le recensement de la population résidant en France est annuel ; tous les habitants ne sont plus recensés la même année. Les communes de moins de

10 000 habitants réalisent une enquête de recensement exhaustive tous les 5 ans, à raison d'un cinquième des communes chaque année. Les communes de 10 000 habitants ou plus réalisent tous les ans une enquête par sondage, auprès d'un échantillon de 8 % des logements. En cumulant cinq enquêtes, l'ensemble des habitants des communes de moins de 10 000 habitants et 40 % environ de la population des communes de 10 000 habitants ou plus sont pris en compte. Les résultats millésimés 2006, date du milieu de la période, correspondent aux cinq premières enquêtes de recensement réalisées de 2004 à 2008. Depuis lors et chaque année, les résultats de recensement sont produits à partir des cinq enquêtes annuelles les plus récentes. Les résultats du recensement de la population sont présentés en utilisant l'âge en années révolues.

Pour l'essentiel de la population, la date de référence est proche du début du mois de janvier. De ce fait, les effectifs totaux d'une génération d'un âge en années révolues donné sont proches des effectifs ayant le même âge atteint l'année précédant l'enquête de recensement. Ainsi, par exemple, le nombre de personnes ayant 39 ans en âge en années révolues au recensement est proche du nombre de personnes ayant 40 ans en âge atteint dans l'année. On pourra donc comparer les effectifs ayant l'âge *a* au recensement de 2006 avec les personnes ayant l'âge *a* +1 aux recensements de 1999 ou antérieurs.

B.1.b. Les populations départementales

Pour les années 1990 et 1999, les estimations de population s'appuient sur les dénombrements issus des deux recensements de la population de mars 1990 et 1999 dont les données sont ramenées au 1er janvier.

Pour les années 2006 à 2009, les estimations de population proviennent du nouveau recensement.

Entre les recensements, la population au niveau départemental est d'abord estimée par l'INSEE de façon globale. Les statistiques de l'état civil fournissent les soldes naturels départementaux et régionaux. Concernant les soldes migratoires, les recensements de population passés permettent de calculer un solde migratoire apparent, obtenu en

soustrayant le solde naturel à la variation de population. Les tendances observées sont ensuite prolongées. Un calage est enfin réalisé afin que la population de l'ensemble des départements ou régions coïncide avec l'estimation de population nationale. Dans un second temps, la population au niveau départemental et régional est répartie par âge, en utilisant les statistiques d'état civil par sexe et âge et en prolongeant les tendances observées sur les flux migratoires par sexe et âge avec les recensements de population. L'âge considéré est l'âge en années révolues au 1er janvier de l'année.

B.2. Estimations du nombre de cas attendus par commune

Entre 1990 et 2010, le découpage communal de la France a évolué suite à des fusions, rattachements de communes ou créations et rétablissements. Pour les communes qui ont fusionné, le code INSEE de la commune après le rattachement leur a été affecté pour toute la période 1990-2010 et les populations de ces communes avant fusion ont été additionnées. Aux communes qui sont apparues ou qui ont été rétablies sur la période a été affecté le code INSEE avant le rétablissement. La population de ces communes a également été additionnée à celle de la commune avant rattachement. Nous avons ainsi divisé le territoire métropolitain en 36 326 entités stables sur toute la période 1990-2010, qui coïncident dans la vaste majorité des cas avec la notion de commune et que nous conviendrons donc par simplification d'appeler commune, même s'il s'agit, dans un faible nombre de cas, de fusions de communes.

Estimation du nombre de personnes-années (PA) à risque de LA par commune

Les populations communales par année, année d'âge et sexe sont disponibles grâce aux recensements INSEE pour les années 1990,1999 et 2006 à 2010 (au 31 décembre 2013). Pour les années 1991 à 1998 (resp. 2000 à 2005), les populations communales par année, année d'âge et sexe ont été estimées à partir des données de population par département, par année, année d'âge et sexe (entre 0 et 15 ans) et de la proportion de population des communes par rapport à celles des départements (recensements INSEE 1990,1999 et 2006).

En faisant l'hypothèse que pour un âge donné, la part, notée *p* de la population d'une commune dans la population d'un département a varié linéairement avec l'année entre les deux recensements, la population des communes par année, année d'âge et sexe a été estimée par interpolation linéaire de cette part pour les années 1991 à 1998 avec 1≤N≤8 :

$$\mathrm{pop}(90+\mathrm{N})_{\mathrm{c,a,s}}=\left(\mathrm{p}(90)_{\mathrm{c,a,s}}+\mathrm{N}\left(\frac{\mathrm{p}(99)_{\mathrm{c,a,s}}-\mathrm{p}(90)_{\mathrm{c,a,s}}}{9}\right)\right)*\mathrm{pop}(90+\mathrm{N})_{\mathrm{dept(c),a,s}}$$

De même pour les années 2000 à 2005 avec 1≤N≤6 :

$$\mathrm{pop}(99+\mathrm{N})_{\mathrm{c,a,s}}=\left(\mathrm{p}(99)_{\mathrm{c,a,s}}+\mathrm{N}\left(\frac{\mathrm{p}(2006)_{\mathrm{c,a,s}}-\mathrm{p}(99)_{\mathrm{c,a,s}}}{7}\right)\right)*\mathrm{pop}(99+\mathrm{N})_{\mathrm{dept(c),a,s}}$$

Où pop(90+N) (resp. pop(99+N)) est l'effectif estimé des sujets l'année 1990+N (resp. 1999+N), *c* la commune, *a* l'âge, *s* le sexe, *dept(c)* le département auquel appartient la commune *c*, $\mathrm{p}(90)_{\mathrm{c,a,s}}$, $\mathrm{p}(99)_{\mathrm{c,a,s}}$ et $\mathrm{p}(2006)_{\mathrm{c,a,s}}$ sont les proportions de population d'âge *a* et de sexe *s* en 1990, 1999 et 2006 de la commune *c* par rapport à celle du département.

Sous l'hypothèse d'une répartition uniforme des naissances au cours de l'année, les personnes-années à risque de LA, notées PA, ont été estimées pour l'année 1990+N, (0≤N≤19) de la façon classique suivante :

$$\mathrm{PA}(1990+\mathrm{N})_{\mathrm{c,a,s}}=\frac{\mathrm{pop}(1990+\mathrm{N})_{\mathrm{c,a,s}}+\mathrm{pop}(1990+\mathrm{N}+1)_{\mathrm{c,a,s}}}{2}$$

Estimation des taux d'incidence de référence des leucémies de l'enfant

Les taux d'incidence (TI) de référence des LA par année d'âge *a* et sexe *s* utilisés sont les taux observés sur l'ensemble de la métropole et ont été estimés par la formule suivante, avec 1≤N≤19 et 1≤c≤36 326 :

$$\mathrm{TIref}(1990+\mathrm{N})_{\mathrm{a,s}}=\frac{\sum_{\mathrm{c}}\mathrm{O}_{\mathrm{c,a,s}}}{\sum_{\mathrm{c}}\mathrm{PA}(1990+\mathrm{N})_{\mathrm{c,a,s}}}$$

où $\mathrm{O}_{\mathrm{c,a,s}}$ est le nombre de cas de LA d'âge *a* et de sexe *s* diagnostiqués l'année 1990+N dans la commune *c*.

Le nombre de cas de LA attendus E dans la commune *c* pour l'année 1990+N, l'âge *a* et le sexe *s* est alors donné par :

$$E(1990+N)_{c,a,s}=TIref(1990+N)_{a,s}*PA(1990+N)_{c,a,s}$$

III. Témoins

Afin de servir de groupe de référence aux cas contemporains pour la période 2002-2007 et dans le cadre de l'étude cas-témoins GEOCAP (GEOlocalisation des CAncers Pédiatriques), un échantillon de 5 000 sujets par année, représentatif de la population française des sujets de moins de 15 ans, en termes d'âge et de répartition spatiale sur la France métropolitaine, a été constitué en collaboration avec l'INSEE, à partir des bases de données de la Direction Générale des Impôts (DGI) sur les déclarations de revenus et taxes d'habitation qui comportent chacune l'adresse du foyer et la liste ses années de naissance des personnes vivant au foyer. Pour chacune de ces années 2002 à 2007, un échantillon de 5 000 enfants a été extrait de la base totale des enfants habitant en France métropolitaine. Afin d'avoir toutes les variables nécessaires à l'échantillonnage et à l'étude ultérieure, l'INSEE a croisé les fichiers d'impôts sur le revenu et de taxe d'habitation. Plus exactement, ce sont les fichiers de l'année N de l'impôt sur le revenu et ceux de l'année N+1 de la taxe d'habitation qui, une fois recoupés, permettent de faire le tirage de l'échantillon correspondant à l'année N. Le tirage de l'échantillon a été fait par la procédure « surveyselect » de SAS, à partir de la base des enfants français de moins de quinze ans, par tirage au sort aléatoire des unités statistiques que sont les membres des foyers fiscaux âgés de moins de 15 ans, après stratification sur le département. Le nombre d'enfants tirés au sort dans chaque département est ainsi proportionnel à la taille de la population française de moins de quinze ans dans ceux-ci. Le tirage a été stratifié par département en fonction de la taille de la population des enfants de moins de quinze ans dans chacun, essentiellement pour faciliter la manipulation de ces fichiers qui sont extrêmement volumineux.

L'échantillon constitué et la population source ont été comparés :

- selon les caractéristiques suivantes du foyer fiscal : âge atteint au 31 décembre, nombre d'enfants du ménage, revenu du foyer fiscal (disponible par foyer « individuel » et utilisé par l'INSEE pour éprouver la représentativité des échantillons annuels mais non transmis),

- et selon des caractéristiques de la commune d'habitation : tranche d'unité urbaine, revenu médian, proportion de HLM, proportions de résidents propriétaires, nombre moyen de personnes par pièce, proportions de bacheliers dans la population de plus de quinze ans, proportion d'ouvriers dans la population active et taux de chômage (Sermage-Faure et al., 2012).

IV. Géolocalisation

Le géocodage consiste à affecter des coordonnées géographiques (longitude/latitude) à une adresse. Ce procédé est fondé sur la mise en œuvre de traitements automatisés de manière ponctuelle ou sur des fichiers d'adresses (individus, entreprises, points d'intérêt, etc.). Les coordonnées géographiques permettent de positionner chaque adresse sur une carte numérique via un Système d'Informations Géographiques (SIG). Le géocodage est une des techniques de géolocalisation (ou « géoréferencement ») définie comme un procédé permettant de positionner un objet (une personne, etc) sur un plan ou une carte à l'aide de ses coordonnées géographiques.

Le système de coordonnées utilisé pour cette étude est le système de coordonnées projetées « Lambert II étendu » qui est un système français de projection permettant de définir des coordonnées (X, Y) sur un même plan pour tout point en France.

A. Géocodage des adresses

Deux ensembles d'adresses ont été géocodés :

- les adresses des mesures de radiations naturelles de la campagne IRSN dans le cadre de l'élaboration d'un indicateur d'exposition à ces radiations (cf partie V. II),

- celles des cas et des témoins de l'étude GEOCAP à la fois dans le cadre de cette élaboration d'indicateurs d'exposition et pour en étudier l'association avec les LA.

Les adresses des domiciles où ont été effectuées les mesures d'activité volumique du radon ont été géocodées. Les adresses au diagnostic des cas et l'adresse du domicile des témoins de l'année considérée ont été géocodées dans l'étude GEOCAP.

En étroite collaboration avec notre équipe, le géocodage a été effectué par l'entreprise Géocible au moyen du logiciel de système d'information géographique MAPINFO (Mapinfo, Pitney Bowes Software Inc., Troy, NY, USA) et des cartes vectorisées NAVTEQ (Navteq, Paris, France; http://www.navteq.com/) et BD ADRESSE (données de l'Institut Géographique National - IGN). Les algorithmes mis en place par Géocible aboutissent, quand c'est possible, à une localisation fournie sous forme de coordonnées Lambert II étendu (X, Y). L'algorithme de géolocalisation permet de corriger de manière automatique les erreurs d'orthographe ou autres erreurs minimes qui peuvent exister dans les adresses fournies en entrée. Cependant, il arrive que malgré ces corrections, les adresses ne soient pas reconnues dans les bases de données. Un travail manuel est alors effectué pour tenter de trouver la localisation de l'adresse fournie : vérification manuelle de l'adresse et/ou utilisation des plans des communes.

B. Précision du géocodage

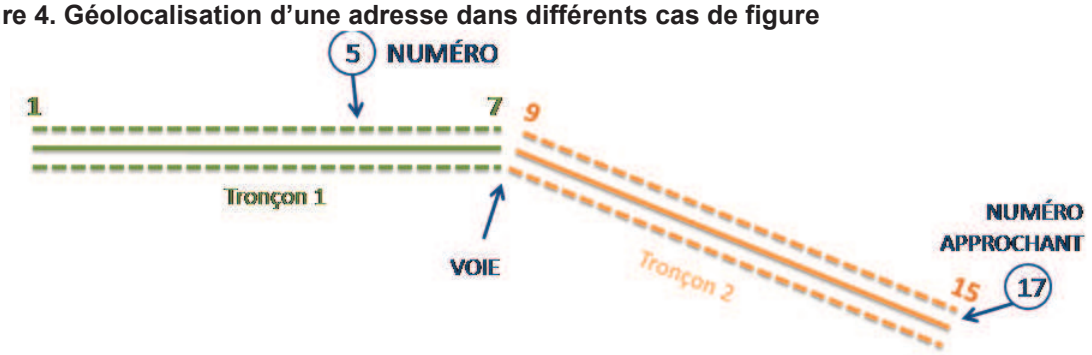
La précision du géocodage est essentielle pour bien estimer l'exposition des sujets. Elle dépend de deux séries d'éléments :

• La qualité des adresses d'entrée qui est de qualité variable ; par exemple : adresses sans numéro s'il n'existe pas ou s'il n'est pas essentiel en utilisation courante et donc non précisé dans les données, lieu-dit renseigné comme seule adresse ne permettant pas une géolocalisation précise. Certains renseignements peuvent aussi être tout simplement manquants.

• La précision du géocodage dépend également de l'information disponible dans les bases de données utilisées. Dans ces bases de données, une voie est découpée en

tronçons dont seules les coordonnées des extrémités sont géolocalisées et pour lesquelles les coordonnées des adresses intermédiaires doivent être estimées si nécessaire (Figure 4).

Figure 4. Géolocalisation d'une adresse dans différents cas de figure



En vert :géolocalisation du numéro 5 d'un tronçon de voie par interpolation linéaire entre les numéros géolocalisés (numéros 1 et 7) des extémités de ce tronçon.
En rouge : géolocalisation d'un numéro non compris dans un tronçon de voie (numéro 17) au numéro géolocalisé le plus proche (numéro 15).

Différents niveaux de précision de localisation des adresses sont ainsi obtenus. La meilleure précision est obtenue quand l'adresse est complète (la commune, la voie et le numéro dans la voie sont fournis) et qu'elle est identifiable dans une des bases de données utilisées : la localisation de l'adresse est alors interpolée entre les coordonnées exactes des extrémités du tronçon l'encadrant. La précision correspondante est « NUMÉRO » (Figure 4, exemple du numéro 5). Quand l'adresse n'est pas complète ou imprécise ou que les bases de données sont moins précisément renseignées pour cette adresse, la localisation est nécessairement moins précise. Quand le numéro précis ne fait pas partie des bases, les numéros approchants (les numéros dans les voies sont souvent indiqués aux croisements avec d'autres voies) peuvent être utilisés pour approcher au mieux la position du numéro de l'adresse et la précision indiquée est alors « NUMÉRO APPROCHANT» (Figure 4, exemple du numéro 17). Lorsque la voie est fournie sans numéro ou qu'aucun numéro n'est donné pour cette voie dans les bases de données, le géocodage est fait au centre de la voie. La précision indiquée « VOIE » (Figure 4) dépend alors de la longueur de la voie, qui est parfois renseignée dans les bases de données. Si l'adresse fournie en entrée est un lieu-dit sans indication de voie, le géocodage est fait au centre du lieu-dit. La précision indiquée est « LIEU-DIT ». Enfin, si seule la commune est fournie en adresse ou si la voie fournie n'est

trouvée ni dans les bases de données ni manuellement, la localisation est faite à la mairie de la commune de résidence : précision « COMMUNE ».

Le Tableau suivant indique la précision obtenue pour le géocodage des deux ensembles d'adresses considérés.

Tableau 1. Type de géocodage obtenu pour les sujets GEOCAP et les mesures de radiations IRSN

Géocodage obtenu	GEOCAP				Mesures de radiations (IRSN)	
	Témoins		Cas			
	N	%	N	%	N	%
NUMERO	23 171	77,2%	1 932	69,9%	5 222	48,2%
NUMERO APPROCHANT	802	2,7%	130	4,7%	104	1,0%
VOIE	4 405	14,7%	463	16,8%	1 982	18,3%
LIEU-DIT	1 330	4,4%	151	5,5%	391	3,6%
COMMUNE	292	1,0%	87	3,1%	3 144	29,0%
TOTAL	30 000		2 763		10 843	

La précision du géocodage est légèrement meilleure chez les témoins, issus des bases de données de la Direction Générale des Impôts, que chez les cas issus du RNHE. Cet écart de qualité reflète un écart de qualité des adresses recueillies entre les cas et les témoins. Il est cependant limité et nous verrons qu'étant donné que nos variables d'exposition seront finalement lissées au niveau des communes ou des ZE, cette différence ne jouera pas sur nos conclusions. La précision du géocodage était plus importante pour le travail réalisé en parallèle sur les expositions aux lignes à haute tension et en analysant la sensibilité des résultats selon la catégorie de précision du géocodage, il a été vérifié à toutes les étapes de ce travail que cet écart n'expliquait aucun des résultats.

Les campagnes de mesures d'exposition aux radiations naturelles à l'intérieur ou l'extérieur de domiciles ont été effectuées par l'IRSN sur plusieurs années, département par département, et dans chaque département grâce à l'intervention d'enquêteurs de la DDASS (Direction Départementale de l'Action Sanitaire et Sociale) concernée. De ce fait la précision des adresses recueillies et leur géolocalisation sont de moindre qualité.

Par ailleurs, en collaboration avec la société Géocible, nous avons élaboré et mis en place un système de détection d'erreurs de géocodage afin de pouvoir les corriger. Pour cela

différents critères de fiabilité ont été mis en place (cohérence de la distance entre le point géocodé et la mairie de la commune indiquée dans l'adresse, vérification de l'appartenance du point au polygone de la commune de départ, etc.). S'ils n'étaient pas vérifiés, le géocodage faisait alors l'objet d'une vérification et d'une correction.

C. Validation du géocodage

En parallèle, dans le cadre de l'étude de l'association entre LA et exposition aux champs magnétiques liés à la proximité des lignes à haute tension du projet GEOCAP, les sujets les plus proches des lignes à haute tension ont été localisés de façon plus précise. Le champ magnétique dû aux lignes à haute tension décroît, en effet, très rapidement avec la distance à la ligne et une plus grande précision dans leur géolocalisation est donc nécessaire à proximité immédiate des lignes. Pour ces sujets, la position exacte du bâtiment de l'adresse a été cherchée manuellement (géocodage « au domicile »). Une validation de la géolocalisation de ces sujets a ainsi été permise, ainsi qu'une évaluation de la précision donnée. La distribution des distances entre le géocodage et le géocodage « au domicile » pour les 2 846 sujets concernés est donnée dans le Tableau 2.

Tableau 2. Distribution de la distance en mètres entre géocodage et géocodage « au domicile » en fonction du géocodage

Précision du géocodage	N	p5%	p25%	médiane	p75%	p95%	p99%
NUMERO	2 369	0	6	13	28	91	246
NUMERO APPROCHANT	123	13	23	60	165	1 010	3 628
VOIE	282	15	43	85	166	615	2 312
LIEU-DIT	31	40	56	174	435	695	1 601
COMMUNE	41	123	362	759	2 333	5 759	11 413
TOTAL	2 846	0	7	16	42	214	1 474

px% : $x^{ème}$ percentile de la distribution

Les médianes des distributions par type de géocodage sont respectivement de 13 m, 60 m, 85 m, 174 m et 759 m. Cette distance peut être considérée comme une bonne estimation de la précision du géocodage « automatique ».

V. Méthodes statistiques générales

A. Etude d'incidence

Dans les études d'incidence, le risque relatif de LA a été estimé par le ratio d'incidence standardisé (SIR), défini comme le ratio du nombre de cas observés (O) sur le nombre de cas attendus (E). Lorsque le nombre de cas attendus est petit (<10), l'approximation normale n'est pas satisfaisante ; et l'intervalle de confiance à 95% du SIR a été calculé en utilisant l'approximation de Byar (Breslow and Day, 1987).

Sous l'hypothèse d'une maladie rare et non-infectieuse, les cas O_i d'une unité géographique i peuvent être considérés comme des variables de Poisson indépendantes. O_i~Poisson(E_i) où E_i est le nombre de cas attendus dans l'unité sous l'hypothèse de taux d'incidence par âge et sexe égaux aux taux de référence correspondant pour la France entière. Dans un modèle quantitatif $\log(E(O_i))=\alpha+\beta X$ une augmentation d'une unité de la variable d'exposition X correspond à une variation de exp(β) du risque relatif. Dans un modèle qualitatif, où la variable X a K classes décrites par k-1 fonctions indicatrices X_k en prenant la classe K comme classe de référence, on peut écrire : $\log(E(O_i))=\alpha+\sum_k \beta X(i)$. exp($\beta_k$) donne l'estimation du risque relatif de l'unité géographique i associé à l'exposition X_k par rapport à la référence. Les analyses ont été réalisées en utilisant la procédure GENMOD de SAS®. Pour tester l'hétérogénéité des SIR dans les analyses, le test classique du Khi-2 de Pearson a été utilisé :

$$Q^2=\sum_i \frac{(O_i-E_i)^2}{E_i}$$

Cette statistique suit asymptotiquement une loi du Khi-2 à (K–1) degrés de liberté sous l'hypothèse nulle d'homogénéité.

Le test de Potthoff et Whittinghill, spécifiquement utilisé pour tester l'hétérogénéité des SIR dans le cadre d'une recherche d'hétérogénéité spatiale est détaillé dans la partie suivante (Partie IV. III. A).

B. Etude cas-témoins

Dans les études cas-témoins, le risque relatif de LA a été estimé par l'odds ratio (OR) calculé par régression logistique non conditionnelle ajusté sur l'âge en classes de un an.

La seule information disponible dans les fichiers de l'INSEE sur l'âge des témoins est leur année de naissance : ce sont donc les âges en années révolues au 31 décembre de l'année du diagnostic pour les cas et à la fin de l'année d'échantillonnage pour les témoins qui ont été utilisés. Pour conserver une comparabilité des groupes de cas et de témoins, les cas du RNHE diagnostiqués à l'âge de 14 ans mais atteignant l'âge de 15 ans au 31 décembre de l'année ont été exclus de l'analyse.

Les analyses statistiques ont été réalisées en utilisant le logiciel SAS software package (version 9.3 ; SAS Institue, Cary, NC).

VI. Discussion

Au lieu de considérer les effectifs pour un âge donné sur les deux recensements, il est aussi possible de prendre en compte une cohorte de naissance que l'on suit de 1990 à 1999 (méthode « diagonale simple »). La différence observée entre les deux recensements pour les sujets d'âge *i* en 1990 et les sujets d'âge *i*+9 en 1999 est prise en compte en faisant l'hypothèse que la variation annuelle est constante entre les deux recensements. La méthode diagonale simple est souvent utilisée pour étudier l'incidence d'un événement, notamment la mortalité, dans une population d'adultes, où la mortalité affectant une cohorte de naissance ne peut généralement pas être considérée comme négligeable. Le problème qui se pose avec la méthode diagonale simple est sa validité pour les premiers âges. En effet, en reprenant l'exemple des recensements 1990 et 1999, les effectifs des cohortes d'âges inférieurs à 9 ans en 1999 ne sont pas connus, puisque ces enfants n'étaient pas nés en 1990. Or dans le cas présent, l'exactitude des données pour les jeunes enfants est primordiale et la mortalité de toutes les cohortes étudiées, entre 0 et 14 ans, très faible ; il nous semble donc plus approprié d'utiliser la méthode horizontale.

Le problème majeur de la méthode horizontale est qu'elle ne prend pas en compte les différences d'effectifs des cohortes de naissance. Cette question est importante lorsque la population considérée comprend des classes d'âge d'effectifs très différents souvent en raison de phénomènes exceptionnels tels que des guerres ou des catastrophes naturelles. Les classes d'âges à effectifs très faibles, appelées classes « creuses », « se déplacent » en effet d'année calendaire en année calendaire, or cette méthode considère que pour un âge donné les effectifs restent sensiblement constants. L'effet de la non-prise en compte des classes creuses est important lorsque l'on considère les classes d'âge de personnes vivantes au moment des deux guerres mondiales, en raison des importants déficits de naissances durant ces périodes, ce qui n'est pas le cas dans la présente étude.

Pour estimer l'erreur induite par ce type d'interpolation, les populations du recensement de 1999 ont été comparées aux estimations calculées par interpolation entre les années 1990 et 2006. Les percentiles 10% et 90% de la distribution du rapport population communale estimée/population du recensement sont respectivement 0,83 et 1,32. La médiane du rapport est 1,02. Ces résultats correspondent à une interpolation sur 17 ans. L'erreur attendue pour des interpolations de 10 ans (entre 1990 et 1999) et 8 ans (entre 1999 et 2006) est moindre.

La qualité de la géolocalisation des adresses de notre étude est un élément important de sa validité. Il existe une nette amélioration dans la dernière version de la base de données IGN (BD ADRESSES version 2.1): les coordonnées d'adresses obtenues par interpolation et projection le long d'un tronçon de route d'une plaque adresse (boîte aux lettres) en version précédente, ont maintenant pour coordonnées, les coordonnées de la plaque adresse. Les adresses complètes sont positionnées à la plaque adresse ou à l'entrée de parcelle. L'amélioration de la géolocalisation des sujets GEOCAP à l'aide de cette base est en cours.

Partie IV Hétérogénéité spatiale de l'incidence des leucémies de l'enfant

I. Introduction

L'analyse de la distribution spatiale de l'incidence des LA est essentielle pour la surveillance de la maladie en vue d'une détection précoce des excès localisés (Goodman et al., 2014). Elle peut également contribuer à l'identification de facteurs de risque environnementaux, notamment. En épidémiologie, la recherche de nouveaux facteurs de risque environnementaux peut adopter diverses approches complémentaires mettant en relation des informations sur les expositions et sur l'incidence des LA. Selon le niveau de connaissances préalable, ces études comme celles menées par l'équipe peuvent être écologiques (Evrard et al., 2005 ; Evrard et al., 2006a) ou individuelles, cas-témoins (Amigou et al., 2011) ou de cohorte. Il peut également être intéressant d'examiner s'il existe une variabilité des taux d'incidence sur un territoire à une échelle donnée, et dans l'affirmative, essayer d'identifier une exposition possiblement corrélée à ces variations. L'analyse de l'hétérogénéité spatiale globale de l'incidence d'une maladie et la localisation des excès potentiels locaux peut donc être informative. De telles études ont été menées dans divers pays par plusieurs auteurs et examinées par McNally et Eden (McNally and Eden, 2004). Par la suite, des études ont été menées en Grande-Bretagne (McNally et al., 2006; McNally et al., 2009), Nouvelle-Zélande (Dockerty et al., 1999a), Palestine (Bailony et al., 2011), Etats-Unis (Amin et al., 2010), Allemagne (Schmiedel et al., 2010) et Danemark (Schmiedel et al., 2011), Hongrie (Nyari et al., 2013). En France, une hétérogénéité modérée de l'incidence des LA sur la période 1990-1994 a été observée à l'échelle des communes avec une surdispersion extrapoissonnienne égale à 1,2 % ($p = 0,06$) (Bellec et al., 2006) et à l'échelle des départements sur la période 1990-2004 (surdispersion de 18,6 %, $p = 0,10$)(Faure et al., 2009). Nous étudions ici la répartition spatiale de l'incidence des LA en France sur la période 1990-2006, à l'échelle géographique intermédiaire des *bassins de vie* (BV), notion définie par l'INSEE comme un territoire constitué de communes et dans lequel la population a accès aux principaux équipements et à l'emploi. Une étude de l'hétérogénéité spatiale globale de la

répartition de l'incidence des LA (clustering) ainsi qu'une recherche spatiale et spatio-temporelle d'excès localisés de LA (clusters) ont été réalisées.

II. Matériel

A. Les cas

Sur la période 1990-2006, 7 675 cas de LA ont été diagnostiqués en France dont 6 269 cas de LAL et 1 276 cas de LAM.

B. Population

En moyenne, un BV est constitué d'un peu plus de 19 communes ce qui représente un peu plus de 30 500 habitants et en moyenne 5 450 enfants (Tableau 3).

Tableau 3. Distribution statistique des bassins de vie français, en nombre de communes, taille de population et effectifs attendus de cas de leucémies de l'enfant.

	Moyenne	Min	Q1[a]	Médiane[a]	Q3[a]	Max
Nb communes par BV	19,1	1	7	13	24	556
Population tous âges (1999)	30 542,0	270,0	6 219,8	9 754,5	17 968,0	9 802 327,0
Population 0-14 ans (1999)	5 453,7	71,0	1 068,0	1 725,0	3 298,0	1 823 195,0
Nombre de cas de LA attendus	4,1	0,1	0,8	1,3	2,4	1 388,3

[a] Les percentiles des distributions présentées (Q1 premier quartile, Médiane, Q3 3ème quartile) sont déterminés en pourcentage du nombre de communes, ou pourcentage cumulé des populations impliquées, ou pourcentage cumulé des nombres de cas attendus.

Pour pouvoir appliquer les méthodes de détection de clusters, les petites îles françaises composées d'un seul *bassin de vie*, sans voisin contigu, n'ont pas été incluses dans notre étude. Ainsi, 1 895 *bassins de vie* ont été étudiés.

Le nombre de cas attendus a été estimé par année et tranche d'âge ([0 ; 1[, [1 ; 5[, [5 ; 10[, [10 ; 15[). Le nombre moyen de cas de LA de 0-14 ans attendus par BV sous l'hypothèse H_0 d'un taux d'incidence homogène sur toute la France était de 4,1 cas pour la période 1990-2006 (Tableau 3). Le BV contenant Paris avait le nombre de cas attendus maximum avec presque 1 388 cas, tandis que dans 25% des BV, environ, le nombre attendu était inférieur à 0,8.

III. Méthodes

A. Méthodes de détection d'une hétérogénéité globale (clustering)

A.1. Test de Potthoff et Whittinghill

La présence d'une sur-dispersion extra-poissonienne a été testée par le test de Potthoff et Whittinghill (Potthoff and Whittinghill, 1966) en postulant que le nombre de sujets ou de cas observés dans un BV est distribué indépendamment pour chacun selon une loi binomiale négative dont le rapport entre sa variance et sa moyenne est égal 1+β. Ainsi, la sur-dispersion est quantifiée par le coefficient d'hétérogénéité β : β=0 correspond à une absence de sur-dispersion (loi de Poisson) alors que β>0 indique que la variance est supérieure à la moyenne (sur-dispersion). Sous l'hypothèse d'homogénéité, la statistique de test proposée par Potthoff et Whittinghill suit asymptotiquement une loi Normale et il a été prouvé que ce test était localement plus puissant que le test du Khi-2 de Pearson, sous cette hypothèse alternative de sur-dispersion poissonienne. Muirhead et Butland (Muirhead and Butland, 1996) ont proposé une estimation du coefficient d'hétérogénéité β et de son écart-type SE(β) sous l'hypothèse nulle avec o_i (resp. e_i) le nombre de cas observés (resp. attendus) dans le BV i, O le nombre total de cas observés dans les 1 895 BV:

$$\hat{\beta}=\frac{\sum_{1\,895}\left(\frac{o_i(o_i\text{-}1)}{e_i}\right)\text{-}(0\text{-}1)}{\frac{(1\,895\text{-}1)(0\text{-}1)}{0}} \text{ et } \mathrm{SE}(\widehat{\beta})=\frac{1}{2}\sqrt{\frac{(1\,895\text{-}1)(0\text{-}1)}{0}}$$

A.2. Test de Rogerson

La statistique R proposée par Rogerson (Rogerson, 1999), est une statistique globale, égale à zéro sous l'hypothèse nulle d'homogénéité spatiale. Cette statistique, basée sur l'indice de Tango, est une combinaison d'un terme d'auto-corrélation entre les unités voisines, semblable à la statistique I de Moran (Moran, 1948) et de la statistique du Khi-2 d'ajustement qui compare les nombres de cas observés et attendus.

La statistique de test est :

$$R=\sum_{i=1}^{1\,895}\frac{\left(\frac{o_i}{O}-\frac{e_i}{E}\right)^2}{\frac{e_i}{E}}+w_{i,j}\sum_{i=1}^{1\,895}\sum_{j=1}^{1\,895}\frac{\left(\frac{o_i}{O}-\frac{e_i}{E}\right)\left(\frac{o_i}{O}-\frac{e_j}{E}\right)}{\sqrt{\frac{e_i}{E}\,\frac{e_j}{E}}}$$

Où o_i (resp e_i) représente le nombre de cas observés (resp attendus) dans le BV i, O (resp E) le nombre total de cas observés (resp attendus) dans les 1 895 BV et w_{ij}=1 si les BV sont voisins (i.e. s'ils ont au moins une frontière commune) et w_{ij}=0 sinon. Nous avons considéré les BV situés dans un rayon de 40 km comme voisins. Des analyses de sensibilité ont été effectuées pour des voisinages de 20 km, 30 km, 50 km et 60 km.

Significativité des tests

La p-value des tests a été évaluée à l'aide de 999 simulations de Monte Carlo. Pour chaque simulation, le nombre total de cas O est réparti dans les BV selon une loi Multinomiale. La statistique de test est calculée pour chaque simulation ce qui donne la distribution de la statistique sous H_0. La p-value est alors:

$$\text{p-value}=\frac{\sum_{t=1}^{T}\{S_t\geq S_{obs}\}}{T+1}$$

où T est le nombre de simulations sous H_0 (T=999), S_t est la statistique de test calculée pour la t^{ème} simulation et S_{obs} la statistique de test observée.

Les analyses de clustering ont été effectuées pour la période 1990-2006 chez les 0-14 ans et pour la tranche d'âge 1-4 ans chez les LAL.

B. Méthodes de détection de clusters

La détection d'agrégats localisés (clusters) est à l'origine du développement de nombreuses méthodes statistiques dont l'une, très répandue, développée par M.Kulldorff et N.Nagarwalla en 1995 (notée SaTScan dans la suite) et d'autres, plus récentes, comme celles développées par T. Tango et K. Takahashi en 2005 (Tango and Takahashi, 2005) (notée FleXScan dans la suite) ou L. Duczmal en 2007 (Duczmal et al., 2007) et Costa en 2012 (Costa et al., 2012). Tandis que la première méthode s'appuie sur des fenêtres mobiles

de forme circulaire ou elliptique, les méthodes plus récentes sont basées sur des algorithmes complexes qui permettent de détecter des clusters de forme plus irrégulière. Avant d'effectuer une recherche de cluster de LA en France, la puissance de ces méthodes a été évaluée.

B.1. Puissance des méthodes

Pour tester la capacité de détection de cluster des méthodes citées ci-dessus, différentes alternatives pour lesquelles il existe un unique cluster ont été envisagées (Figure 5):

- De différentes formes : irrégulière (linéaire ou en U) ou compacte (forme intermédiaire entre une ellipse et un cercle)
- Dans différents lieux (avec un nombre de cas attendus plus ou moins élevé : 20, 45 et 115) : Sarthe, Gironde, Bourgogne
- Avec un risque relatif (RR) variable : 1,5 ; 2 ; 3

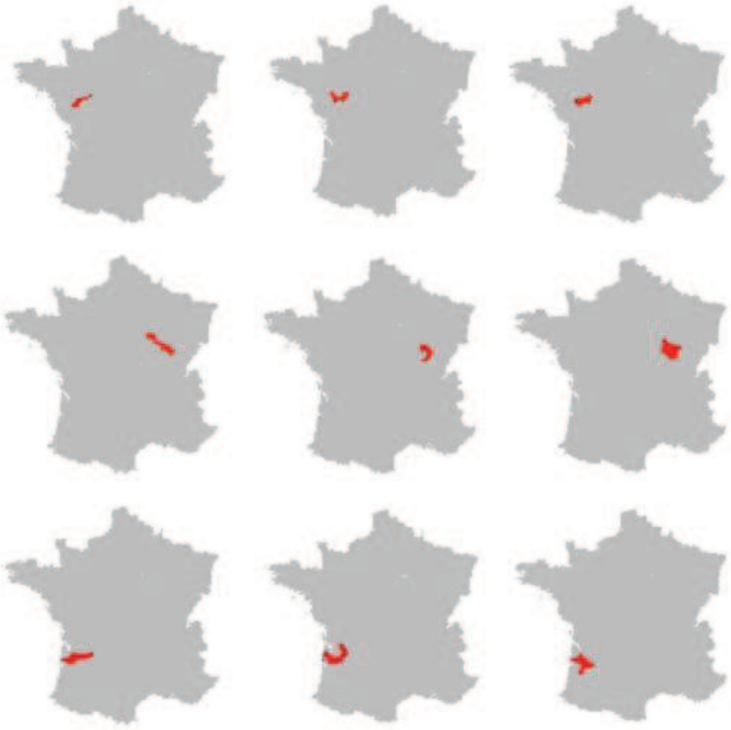
Les différentes hypothèses alternatives envisagées ont été simulées sous l'hypothèse H_1 suivante :

$$o_i \sim \text{Poisson}(\theta_i e_i) \text{où } \theta_i = \begin{cases} \text{RR si i } \epsilon \text{ cluster} \\ 1 \text{ si i } \notin \text{ cluster} \end{cases}$$

et o_i et e_i les nombres de cas observés et attendus dans le BV i.

Pour cela, un nombre total de cas O a été réparti dans les 1 895 BV selon une loi Multinomiale : $M\left(O, \frac{\theta_i e_i}{\sum_{1895} \theta_i e_i}\right)$. Sous cette distribution, le nombre de cas attendus dans le BV i n'est pas exactement égal à $\theta_i E_i$ $(\text{i.e } O * \frac{\theta_i e_i}{\sum_{1895} \theta_i e_i})$ mais le rapport entre les nombres de cas attendus dans les unités géographiques appartenant au cluster et dans les autres est bien égal à RR fois le rapport des nombres de cas attendus correspondant sous H_0.

Figure 5. Hypothèses de répartition non-uniforme des taux d'incidence des LA considérées dans cette étude



Dans chacune des 9 cartes, le taux d'incidence des LA est supposé constant dans la partie grise et plus élevé dans la partie rouge (« cluster »). Les 3 colonnes de cartes diffèrent par la forme des clusters (linéaire, en U, compacte) et les 3 lignes par la localisation (dans la Sarthe, la Bourgogne et la Gironde) et la taille (20, 45 et 115 cas attendus) du cluster.

250 simulations ont été réalisées pour chaque alternative pour estimer une puissance de 80% avec une marge d'incertitude statistique maximale de ± 5%. Pour chaque méthode, une base de 10 000 simulations a été réalisée sous l'hypothèse H_0 d'uniformité du taux d'incidence des LA sur tout le territoire. La p-value pour le cluster le plus probable observé a été estimée par le rapport du nombre de simulations pour lesquelles la statistique de test était supérieure ou égale à la valeur observée au nombre total de simulations sous H_0 (10 000). Pour privilégier la puissance du test de détection d'un cluster par rapport au risque d'inclure un plus grand nombre de faux positifs, nous avons fixé le seuil alpha à 10%.

Le périmètre d'un cluster détecté comme significatif par le test étudié peut ne pas coïncider du tout, ou seulement partiellement avec celui du « vrai cluster » sous-jacent. De ce fait, nous avons considéré d'une part, la probabilité de détecter un cluster, qu'il soit bien ou mal placé, d'autre part, la probabilité de détecter un cluster considéré comme mieux

placé, estimée par la proportion de simulations pour lesquelles le test de détection de cluster est significatif avec au moins un BV vrai positif, c'est-à-dire un BV du « vrai cluster » qui est correctement détecté. La sensibilité et la spécificité des méthodes ont aussi été étudiées.

Cette étude (Goujon-Bellec, Demoury et al., 2011) a montré qu'aucune méthode n'était assez puissante pour détecter et localiser des clusters de petite taille (E ≤45) et de risque relatif faible (RR = 1,5). Cependant, dans des situations de meilleure puissance, les méthodes diffèrent dans leur capacité à localiser le véritable cluster : la fenêtre elliptique de la méthode SaTScan est plus performante dans la détection de clusters linéaires et compacts, et la méthode FleXScan l'est pour les clusters en forme de U. Ces deux méthodes ont été retenues pour la détection de cluster sur le territoire français pour la période 1990-2006.

B.2. Méthodes retenues

B.2.a. Démarche commune aux deux méthodes

Les méthodes de détection de clusters présentées ci-dessous s'inspirent toutes de la démarche proposée par Kulldorff et Nagarwalla en 1995 (Kulldorff and Nagarwalla, 1995).

Dans un premier temps, un ensemble de zones z de clusters potentiels est créé, chaque zone étant composée de BV voisins. La manière dont cet ensemble de zones est déterminé par chacune des méthodes sera détaillée dans les paragraphes suivants.

Soient p et q les SIR de maladie respectivement à l'intérieur et à l'extérieur d'une zone z. Sous l'hypothèse nulle H_0 d'homogénéité des taux d'incidence par âge de la maladie sur tout le territoire, les SIR de maladie à l'intérieur et à l'extérieur de n'importe quelle zone sont égaux (p=q). Sous l'hypothèse alternative H_1, que z est une zone de cluster, le SIR de maladie est supérieur à l'intérieur de z (p>q).

Si le nombre de cas observés dans une zone z suit une distribution de Poisson, il est possible de calculer, pour z, le rapport des vraisemblances suivant:

$$\mathrm{LR}(z)=\frac{L_1(z)}{L_0}=\left(\frac{O_z}{e_z}\right)^{o_z}\left(\frac{O-o_z}{E-e_z}\right)^{O-o_z}1\left\{\frac{o_z}{e_z}>\frac{O-o_z}{E-e_z}\right\}$$

où $L_1(z)$ et L_0 sont les vraisemblances sous les hypothèses H_1 et H_0; o_z et e_z représentent, respectivement, les nombres de cas observés et attendus dans la zone z sur toute la période d'étude ; O et E représentent les nombres de cas observés et attendus sur cette période sur tout le territoire. 1{...} est la fonction indicatrice qui vaut 1 lorsque le rapport des nombres de cas observés et attendus est plus grand dans la zone z que sur le reste du territoire, 0 sinon.

La zone de « cluster le plus probable » (CPP) est ensuite définie comme la zone z qui maximise le rapport LR(z) et la statistique de test est égale à $\max_z(LR(z))$.

Pour chaque méthode, la significativité d'un CPP a été évaluée à l'aide de simulations de Monte Carlo sous l'hypothèse H_0 que le taux d'incidence des cas est homogène sur tout le territoire.

Ces simulations ont permis d'obtenir un échantillon de valeurs issues de la distribution de la statistique de test sous H_0 et la p-value pour le cluster observé a été estimée par le rapport du nombre de simulations pour lesquelles la statistique de test était supérieure ou égale à la valeur observée au nombre total de simulations effectuées sous H_0.

B.2.b. Détermination des zones de clusters potentiels et spécificité des méthodes

Dans ce paragraphe est présenté, pour chaque méthode, l'algorithme qui détermine l'ensemble de zones z de clusters potentiels, à partir duquel est calculée la statistique de test, ainsi que certaines modifications qui sont propres à chaque méthode.

SaTScan

Fenêtre mobile circulaire spatiale

M.Kulldorff et N.Nagarwalla (Kulldorff and Nagarwalla, 1995) ont proposé une méthode de détection de clusters basée sur une fenêtre de forme circulaire qui se déplace sur tout le territoire étudié. Cette fenêtre est tour à tour centrée sur chaque BV et grandit de manière continue en incluant à chaque fois les BV dont le centre se trouve dans le cercle. Les cercles ainsi déterminés définissent l'ensemble des zones z de clusters potentiels. Nous avons fixé à

100 BV la taille maximale d'un cercle. SaTScan[2] est le nom du logiciel qui permet de l'appliquer.

Fenêtre mobile elliptique spatiale

En 2006, M. Kulldorff a introduit des fenêtres de forme elliptique pour la construction des zones de clusters potentiels (Kulldorff et al., 2006). Pour chaque cercle défini précédemment, plusieurs ellipses sont aussi envisagées en faisant varier, d'une part, la longueur *s* d'une ellipse, définie comme le rapport entre l'axe le plus long et l'axe le plus court et d'autre part, l'angle θ entre l'axe le plus long et l'horizontale. Pour chaque valeur de *s*, plusieurs angles θ par rapport à l'horizontale sont ainsi envisagés. Pour une ellipse faiblement allongée (le cas extrême étant le cercle pour *s* =1) la zone recouverte varie peu avec l'angle θ. Au contraire, pour une ellipse allongée, un plus grand nombre d'angles θ est nécessaire pour recouvrir tout le voisinage du BV sur lequel est centrée l'ellipse. C'est pourquoi les auteurs conseillent d'étudier un nombre d'angles égal à trois fois la valeur de *s.*

Le voisinage étudié par une ellipse allongée centrée en un BV est très variable selon la valeur de l'angle θ, tandis que la zone couverte par un cercle ou une ellipse compacte varie moins. L'ellipse allongée est donc plus à même de regrouper deux BV où le risque de maladie est élevé même si elles ne sont pas réellement voisines, et ce même sous une configuration H_0 (Kulldorff et al., 2006).

Pour éviter ce phénomène, il est conseillé d'une part, de limiter l'étude à des ellipses pour lesquelles *s* est compris entre 1 et 5 et d'autre part, d'introduire la fonction de pénalité suivante sur le logarithme du rapport de la vraisemblance LLR(z)=log(LR(z)).

$$\mathrm{LLR}^{\mathrm{pen}}(z)=\mathrm{LLR}(z)*\left(\frac{4s}{(s+1)^2}\right)^{a} \text{ avec } a\geq 0$$

[2]"SaTScanTM is a trademark of Martin Kulldorff. The SaTScanTM software was developed under the joint auspices of MartinKulldorff, the National Cancer Institute, and FarzadMostashari of the New York City Department of Health and Mental Hygiene."

$LLR^{pen}(z)$ tend vers 0 quand s tend vers l'infini, donc plus l'ellipse est allongée, plus le logarithme de la vraisemblance de la zone z qui lui est associé est pénalisé. Le paramètre *a* permet de contrôler la force de la pénalité. Nous avons réalisé l'analyse pour des pénalités nulle (*a*=0) et forte (*a*=1).

Pour cette méthode basée sur une fenêtre mobile elliptique, il est donc possible de considérer comme statistique de test le maximum du rapport LLR(z) ou le maximum du rapport pénalisé $LLR^{pen}(z)$.

Fenêtre mobile spatio-temporelle

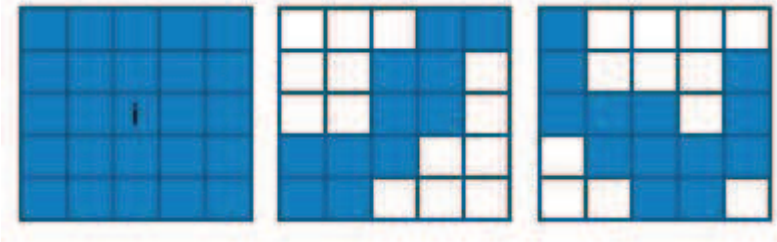
La fenêtre spatio-temporelle est définie par une fenêtre cylindrique avec une base géographique circulaire (ou elliptique) et une hauteur correspondant au temps. La base est définie exactement comme pour l'analyse purement spatiale tandis que la hauteur reflète les périodes de temps potentielles. La fenêtre cylindrique est alors déplacée dans l'espace et le temps de sorte que pour chaque emplacement géographique et taille possible (fixée ici à 100 BV maximum), chaque période de temps (fixée ici à un an) est également possible.

FleXScan

En 2005, T. Tango et K. Takahashi ont proposé une méthode qui identifie les zones de clusters potentiels à partir d'une fenêtre de forme irrégulière (Tango and Takahashi, 2005). Dans un premier temps, cette fenêtre est centrée sur un BV i et composée des K-1 plus proches voisins de ce BV i (K étant fixé par l'utilisateur). Dans un second temps, toutes les zones incluses dans la fenêtre et composées de BV connectés entre eux sont étudiées, deux BV étant connectés s'ils ont une frontière commune. Finalement, toutes les fenêtres, composées de K BV, ainsi que toutes les zones connectées incluses dans ces fenêtres constituent des zones z de clusters potentiels.

La Figure 6 sur laquelle les BV sont représentés schématiquement par des carrés, présente trois exemples de zones connectées incluses dans la fenêtre composée d'un BV i et de ses 24 voisins, cette fenêtre représentant la première zone possible.

Figure 6. FleXScan : zones de clusters potentiels



Sans contrainte particulière sur les BV qui composent le cluster potentiel, les clusters détectés par la méthode FleXScan peuvent être composés de BV à faible SIR de maladie compris entre des BV à haut SIR. Pour éviter ce phénomène, une restriction permettant de prendre en compte de façon individuelle le SIR de maladie de chaque BV peut être introduite dans la statistique de test qui s'écrit alors (Tango, 2008) :

$$T = \max_z \left\{ \left(\frac{o_z}{e_z}\right)^{o_z} \left(\frac{O-o_z}{E-e_z}\right)^{O-o_z} 1\left\{\frac{o_z}{e_z} > \frac{O-o_z}{E-e_z}\right\} \prod_{i \epsilon Z} I(p_i < \alpha_1) \right\} = \max_z \left\{ LR(z) \prod_{i \epsilon Z} I(p_i < \alpha_1) \right\}$$

Où p_i est la p-value modifiée du test H_0 : $E(O_i)=E_i$ donnée par

$$p_i = P(O_i \geq o_i + 1) + \frac{1}{2} P(O_i = o_i) \text{avec } O_i \sim \text{Poisson}(E_i)$$

La p-value p_i introduite ici est inférieure à la p-value usuelle pour compenser le fait que le test basé sur une distribution de Poisson est trop conservateur lorsque le nombre de cas attendus est petit. Avec cette nouvelle statistique de test, seules les zones composées de BV où le risque de maladie est significatif au niveau α_1 (fixé arbitrairement à 0.20) sont des clusters potentiels. Il est ainsi possible d'appliquer cette méthode en considérant comme statistique de test le maximum du rapport LR(z) ou la statistique T. Cette méthode s'applique grâce au logiciel du même nom[3].

Un nombre de 25 et 100 BV ont été considérés comme taille maximale de cluster, avec pénalité (au seul significatif α de 0,20). Sans pénalité, du fait du temps de calcul, seul un voisinage de 25 BV a été considéré.

[3] Takahashi K, Yokoyama T and Tango T. FleXScan: Software for the Flexible Scan Statistic. National Institute of Public Health, Japan, 2005 - 2006.

Les méthodes de détection de clusters ont été utilisées pour la période 1990-2006 et pour les trois sous-périodes 1990-1994, 1995-2000 et 2001–2006 chez les 0-14 ans (toutes LA, LAL, LAM) et pour la tranche d'âge 1-4 ans chez les LAL. Pour privilégier la puissance de détection d'un vrai cluster par rapport au risque d'inclure un plus grand nombre de faux positifs, nous avons fixé le seuil alpha à 10%.

IV. Résultats

Le test de Potthoff et Whittinghill n'a pas montré de surdispersion statistiquement significative pour les LA (β=-0,04 ; p=0,91), LAL ou LAM. De même, le test de Rogerson n'a pas montré d'hétérogénéité significative des taux d'incidence (LA R=0,25 ; p=0,39) (Tableau 4). Lorsque la limite de voisinage a été fixée à 20, 30, 50 et 60 km au lieu de 40 km, les résultats restaient inchangés.

Tableau 4. Test de l'existence d'une sur-dispersion extra-poissonienne des taux d'incidence des leucémies de l'enfant en France métropolitaine (1990-2006) observés dans les 1 895 bassins de vie métropolitains

	Test de P-W		Test de Rogerson	
	$\hat{\beta}$ [a]	p-value	$\hat{R}$	p-value
LA (n = 7 675)	-0,04	0,91	0,25	0,39
LAM (n = 1 276)	-0,04	0,90	1,66	0,21
LAL (n = 6 269)	-0,03	0,86	0,32	0,30
LAL 1-4 ans (n = 3 013)	-0,05	0,93	0,57	0,74

[a] Si O_i est le nombre de cas observés dans BV_i et E_i le nombre de cas attendus correspondant le paramètre de sur-dispersion ß est défini par $var(O_i) = (1 + ß) E_i$, P-W teste l'écart à zéro de la valeur ß. R teste conjointement l'existence d'une autocorrélation entre nombre de cas observés dans des BV voisins et celle d'une surdispersion extra-poissonnienne.

Sur la période 1990-2006, aucun des CPP identifiés avec les méthodes SaTScan et FleXScan spatiales n'étaient significatifs (Tableau 5).

Aucun cluster significatif n'a été identifié en utilisant la méthode spatio-temporelle de SaTScan (Tableau 6).

Tableau 5. Clusters les plus probables détectés par les méthodes SaTScan et FleXScan spatiales sur la période 1990-2006.

Méthode			SaTScan			FleXScan		
Forme de la fenêtre			Circulaire	Elliptique		Flexible		
				Sans pénalité	Avec pénalité	Sans restriction	Avec restriction	Avec restriction
Taille maximale du cluster			100	100	100	25	25	100
LA		O [a]	263	222	256	46	33	104
		E [b]	212,1	167,2	200,0	23,1	15,5	61,4
		BV [c]	73	42	67	10	5	14
		p-value	0,93	0,74	0,68	0,89	0,80	0,20
LAM		O [a]	14	47	60	15	13	13
		E [b]	4,8	23,8	34,9	3,4	2,9	2,9
		BV [c]	11	71	85	9	6	6
		p-value	0,86	0,50	0,55	0,36	0,31	0,51
LAL	0-14 ans	O [a]	244	250	219	34	26	57
		E [b]	187,9	187,0	163,2	13,3	9,7	30,1
		BV [c]	70	71	67	10	6	10
		p-value	0,23	0,27	0,23	0,35	0,36	0,54
	1-4 ans	O [a]	9	113	9	16	15	15
		E [b]	2,1	76,0	2,1	4,3	4,4	4,4
		BV [c]	6	50	6	12	5	5
		p-value	0,74	0,79	0,88	0,76	0,74	0,86

[a] O : nombre de cas observés dans le CPP; [b] : nombre de cas attendus dans le CPP; [c] nombre de BV dans le CPP

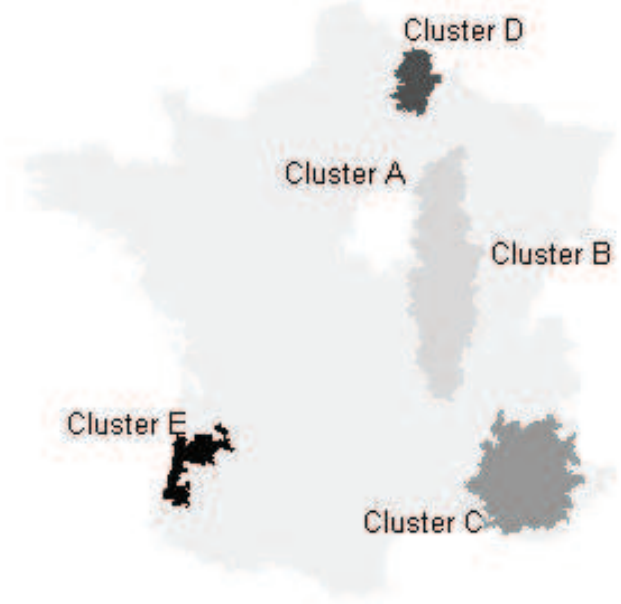
Tableau 6. Clusters les plus probables détectés par la méthode SaTScan spatio-temporelle sur la période 1990-2006

Forme de la fenêtre (avec une taille maximale de 100 BV)			Circulaire	Elliptique	
				Sans pénalité	Avec pénalité
LA		O [a]	52	51	51
		E [b]	25,2	23,0	23,0
		BV [c]	34	28	28
		Période	1990-1995	1990-1995	1990-1995
		p-value	0,57	0,77	0,33
LAM		O [a]	4	8	4
		E [b]	0,1	0,7	0,1
		BV [c]	4	4	3
		Période	2004-2005	1991-1995	2004-2005
		p-value	0,71	0,78	0,81
LAL	0-14 ans	O [a]	122	111	122
		E [b]	76,3	66,4	76,3
		BV [c]	64	67	64
		Période	1995-1999	1997-2003	1995-1999
		p-value	0,36	0,79	0,56
	1-4 ans	O [a]	11	14	14
		E [b]	2,0	2,7	2,7
		BV [c]	13	20	20
		Période	2004-2006	2004-2006	2004-2006
		p-value	0,82	0,96	0,77

[a] O : nombre de cas observés dans le CPP; [b] : nombre de cas attendus dans le CPP; [c] nombre de BV dans le CPP

Pour les trois sous-périodes, quatre CPP détectés par la méthode SaTScan spatiale étaient statistiquement significatifs (Figure 7, Tableau 7).

Figure 7. Clusters spatiaux détectés par les méthodes SaTScan et FleXScan



Pour la période 1990-1994, un cluster de LA a été détecté par la fenêtre elliptique SaTScan avec pénalité (cluster A, p=0,08) mais n'était pas significatif en absence de pénalité. Ce cluster est inclus dans un CPP non significatif identifié par la fenêtre circulaire sur cette période. Le cluster correspond exactement à la base géographique du CPP non significatif trouvé avec la méthode spatio-temporelle de SaTScan pour la période 1990-1995. Pour 1995-2000, un grand cluster de LAM a été détecté avec la fenêtre elliptique SaTScan sans pénalité (cluster B, p=0,08). Il chevauche partiellement les CPP détectés avec les autres formes de fenêtres SaTScan. Un cluster de LAL a été détecté avec les fenêtres circulaire et elliptique avec pénalité (cluster C, p=0,06 et p=0,08, respectivement) avec 24 des 64 BV inclus dans le CPP détecté par la fenêtre elliptique sans pénalité. Le cluster a également été détecté dans l'analyse spatio-temporelle pour la période 1995-1999 mais était non significatif. Un cluster de LAL pour les cas de 1-4 ans a été mis en évidence par la fenêtre elliptique avec pénalité SaTScan (cluster D, p=0,09). Il était non significatif sans pénalité et comprend le CPP identifié par la fenêtre de forme circulaire. La méthode

FleXScan avec une restriction a détecté un cluster de LAL 1-4 ans sur la période 2001-2006 (cluster E, p=0,08) avec la taille maximale de la fenêtre fixée à 25 BV. Ce cluster chevauchait un cluster de 13 BV (p=0,05) détecté avec la taille de fenêtre maximale de 100 BV (non représenté). Le cluster E était inclus dans le CPP non significatif identifié avec la méthode SaTScan spatiale pour la période 2001-2006 et avec la méthode spatio-temporelle pour la période 2004-2006. Comme le montre la Figure 7, les cinq clusters ne se chevauchent pas. Bien que la méthode FleXScan n'ait détecté aucun autre cluster statistiquement significatif, la méthode a identifié des CPP non significatifs dans les zones où la méthode SaTScan avait détecté des clusters (A, B, C et D).

Aucun des principaux clusters identifiés dans une période donnée n’est demeuré statistiquement significatif lorsque la correction de Bonferroni a été appliquée pour tenir compte de la multiplicité des tests (3 périodes de temps considérées et 6 méthodes considérées, soit 18 tests, pour chaque type de LA).

Tableau 7. Clusters les plus probables détectés par les méthodes SaTScan et FleXScan pour les sous-périodes 1990-1994 ; 1995-2000 ; 2001-2006

		CPP [d] identifiés pour la période 1990-1994					CPP [d] identifiés pour la période 1995-2000					CPP [d] identifiés pour la période 2001-2006				
Méthode		SaTScan			FleXScan		SaTScan			FleXScan		SaTScan			FleXScan	
Forme de la fenêtre		Cercle	Ellipse pénalité		Flexible restriction		Cercle	Ellipse pénalité		Flexible restriction		Cercle	Ellipse pénalité		Flexible restriction	
			Non	Oui	Non	Oui		Non	Oui	Non	Oui		Non	Oui	Non	Oui
Taille maximale du cluster		100	100	100	25	25	100	100	100	25	25	100	100	100	25	25
LA	O [a]	43	42	42	26	24	7	14	7	34	8	123	16	54	26	14
	E [b]	21,0	19,1	19,1	8,8	7,9	1,1	3,4	0,8	12,8	1,1	87,0	4,6	30,4	9,4	3,3
	BV [c]	34	28	28	10	7	4	9	3	17	4	79	7	31	11	9
p-value (nom du cluster)		0,10	0,21	**0,08(A)**	0,35	0,17	0,44	0,48	0,24	0,20	0,43	0,54	0,68	0,53	0,60	0,32
LAM	O [a]	8	16	6	8	10	53	41	51	10	12	5	10	5	7	7
	E [b]	1,5	3,9	0,7	1,0	1,9	29,5	18,1	26,6	1,8	3,3	0,4	1,8	0,3	0,8	0,8
	BV [c]	7	21	9	8	8	97	96	96	9	4	4	22	3	5	5
p-value (nom du cluster)		0,46	0,12	0,49	0,41	0,20	0,16	**0,08(B)**	0,11	0,61	0,64	0,20	0,35	0,22	0,47	0,12
LAL 0-14 ans	O [a]	7	81	7	26	23	136	139	136	29	87	111	108	116	21	13
	E [b]	1,1	49,4	1,1	10,1	8,6	91,7	91,4	91,7	10,8	54,0	77,8	69,6	77,5	7,0	3,4
	BV [c]	2	69	2	11	8	64	48	64	16	11	97	70	76	10	5
p-value (nom du cluster)		0,44	0,56	0,54	0,87	0,55	**0,06(C)**	0,11	**0,08(C)**	0,43	0,45	0,78	0,44	0,49	0,82	0,68
LAL 1-4 ans	O [a]	15	7	5	14	14	18	25	25	19	12	62	67	62	17	17
	E [b]	5,0	0,9	0,5	3,5	3,5	5,9	8,7	8,7	4,7	2,6	37,0	38,1	37,0	4,1	4,1
	BV [c]	18	7	5	7	7	20	24	24	13	10	75	76	75	10	10
p-value (nom du cluster)		0,57	0,61	0,63	0,71	0,28	0,22	0,23	**0,09(D)**	0,11	0,32	0,37	0,33	0,55	0,23	**0,08(E)**

[a] O : nombre de cas observés dans le CPP; [b] : nombre de cas attendus dans le CPP; [c] nombre de BV dans le CPP ; [d] CPP : cluster le plus probable

V. Discussion

Cette étude de la répartition spatiale des taux d'incidence de LA, LAL et LAM au cours de la période 1990-2006 à l'échelle des BV n'a pas mis en évidence d'hétérogénéité globale, ni avec le test de Potthoff et Whittinghill (Potthoff and Whittinghill, 1966) ni avec celui de Rogerson (Rogerson, 1999). Comme la recherche de clusters spatio-temporels pour les années 1990-2006, l'investigation de clusters spatiaux pour l'ensemble de la période, n'a pas donné non plus de résultats significatifs, ni avec la méthode SaTScan (Kulldorff and Nagarwalla, 1995), ni avec la méthode FleXScan (Tango and Takahashi, 2005). Cependant, lorsque les trois périodes (1990-1994, 1995-2000 et 2001-2006) ont été considérées, cinq clusters spatiaux statistiquement significatifs ont été détectés. Les clusters spatiaux identifiés dans les analyses par période étaient spécifiques aux différents sous-types de LA, situés dans des zones disjointes et observés pour des périodes différentes (LA période 1990-1994; LAM 1995-2000; LAL 1995-2000; LAL 1-4 ans 1995-2000 et 2001-2006). Dans toutes les situations, les CPP identifiés en utilisant les méthodes SaTScan et FleXScan étaient situés dans la même région, et, à l'exception des LAM, les clusters coïncidaient avec la base géographique des clusters non significatifs identifiés dans l'analyse spatio-temporelle. Les résultats de ces méthodes étaient donc similaires dans l'identification des zones de cluster le plus probable mais pas en ce qui concerne le niveau significativité. Dans les méthodes de détection de cluster, un cluster potentiel est identifié comme cluster le plus probable, parmi toutes les zones candidates, s'il maximise la fonction de rapport de vraisemblance. La p-value tient compte du nombre de zones candidates considérées, ce qui est étroitement lié à la taille de la zone d'étude et à la taille de la fenêtre de balayage (Wheeler, 2007). En outre, la puissance des méthodes de détection de cluster dépend de la répartition spatiale des unités et de la taille de leurs populations, comme cela a été précédemment suggéré (Huang et al., 2008; Song and Kulldorff, 2003; Waller et al., 2006). L'étude de puissance menée sur les données françaises (Goujon-Bellec et al., 2011) afin d'évaluer la performance des méthodes SaTScan, FleXScan et de deux autres méthodes plus récentes (Costa et al.,

2012; Duczmal et al., 2007) a montré que dans les situations où le «vrai» cluster était de moyenne ou de grande taille (soit plus de 45 cas de LA prévus sur la période 1990-2006) et le rapport des risques relatifs supérieur ou égal à deux, toutes les méthodes avaient une puissance de plus de 90%. Les méthodes SaTScan avec fenêtre elliptique et FleXScan avaient la sensibilité la plus élevée. Si un tel cluster existait en France, il aurait donc été nettement détecté par les méthodes SaTScan et FleXScan. L'étude de puissance a également montré que les clusters avec 20 cas attendus et un rapport de risques relatifs égal à 1,5 ou 2 étaient presque indétectables par l'une de ces méthodes. Cette faible puissance peut expliquer la discordance entre les résultats des méthodes SaTScan et FleXScan dans la présente étude, puisque tous les clusters identifiés excepté un étaient composés de moins de 20 cas attendus.

Quelques clusters de maladies rares, telles que les LA de l'enfant, ont été détectés avec les méthodes de détection de clusters. Parmi les neuf études ayant utilisé la méthode SaTScan (Amin et al., 2010; Bailony et al., 2011; Bellec et al., 2006; Hjalmars et al., 1996; Kulldorff and Nagarwalla, 1995; Mosavi-Jarrahi et al., 2007; Schmiedel et al., 2010; Schmiedel et al., 2011; Wheeler, 2007), seulement trois clusters spatiaux ou spatio-temporels ont été détectés (Amin et al., 2010; Kulldorff and Nagarwalla, 1995; Wheeler, 2007). Aucune cause n'a encore été identifiée.

L'échelle et les frontières sont déterminantes dans l'analyse de l'hétérogénéité spatiale. L'échelle géographique à laquelle l'hétérogénéité spatiale est étudiée peut masquer l'hétérogénéité qui peut exister sur une plus grande échelle, telle que l'échelle des départements (Faure et al., 2009) ou entraver la détection d'une hétérogénéité présente à petite échelle, telle que celle des communes (Bellec et al., 2006). Les clusters peuvent évidemment franchir les frontières de recensement et donc être moins susceptibles d'être identifiés. Dans cette étude, les analyses ont été effectuées à l'échelle des BV. Les BV sont des zones d'activité homogènes et sont structurés autour des flux migratoires quotidiens de la population. La population de ces unités non administratives est homogène aux niveaux géographique, social, culturel et économique. Cette échelle semble donc bien appropriée

pour la détection d'excès de cas sur de grands espaces et compatible avec les expositions partagées par la plupart de la population d'un BV.

Parce que les BV sont hétérogènes en nombre d'habitants, la taille maximale des clusters a été spécifiée en termes de nombre de BV et non d'une proportion donnée de la population totale, par exemple 50 %, comme cela se fait souvent. Une fenêtre de taille 10 % de la population aurait été un choix raisonnable. Cependant, avec cette limite, quatre BV, dont le BV centré sur Paris, et toutes les fenêtres centrées sur celui-ci n'auraient pu être considérées comme candidates.

En conclusion, quelques clusters spatiaux ont été identifiés dans des lieux et périodes spécifiques. Bien que les niveaux de significativité des clusters ne soutiennent pas fortement l'existence de facteurs de risque localisés, les clusters peuvent montrer un léger impact de facteurs de risque partagés à l'échelle des BV, y compris dans le contexte d'expositions environnementales. Cependant, cette étude n'a pas mis en évidence d'hétérogénéité spatiale des taux d'incidence des LA de l'enfant sur l'ensemble de la période 1990-2006 et sur l'ensemble de la France métropolitaine.

Il n'en reste pas moins intéressant d'examiner si l'incidence des leucémies est liée à une exposition environnementale particulière, comme le radon, la puissance de détection d'un lien écologique avec une exposition particulière étant nettement supérieure à celle de la détection de l'hétérogénéité globale qu'engendre ce lien spécifique sous-jacent. Notons aussi, par exemple, que tous les clusters significatifs ont été détectés pour des années antérieures à 2001 (clusters A, B, C et D) et sont globalement situés dans l'Est de la France où les retombées radioactives (césium 137) de l'accident de Tchernobyl ont été les plus importantes (Verger P, 2009). Cependant aucun cluster n'est observé dans les Alpes et les régions Alsace et Lorraine, relativement plus affectées par les retombées post accident de Tchernobyl. Si l'observation des clusters A, B, C et D peut suggérer des pistes à explorer, elle ne peut permettre à elle-seule de considérer que ces pistes seraient déjà des hypothèses véritablement « étayées » par des observations.

Partie V Exposition aux radiations ionisantes d'origine naturelle

I. Introduction

A. La radioactivité

La radioactivité est un phénomène physique naturel au cours duquel des noyaux atomiques instables, dits radio-isotopes, se transforment spontanément (« désintégration »), en dégageant de l'énergie sous forme de rayonnements divers, en des noyaux atomiques plus stables ayant perdu une partie de leur masse. Les rayonnements ainsi émis sont appelés rayons alpha, bêta ou gamma. Les deux principales causes d'instabilité sont un trop grand nombre de nucléons ou un déséquilibre entre les nombres de protons et de neutrons. Dans le premier cas de la radioactivité alpha, le noyau évolue vers une plus grande stabilité en émettant une particule alpha. Dans le second cas, un proton se transforme en un neutron (ou l'inverse), avec émission d'un électron ou d'un positon, c'est la radioactivité bêta. La radioactivité gamma est une simple désexcitation du noyau, de même nature que l'émission de lumière ou de rayons X par les atomes. Les désintégrations gamma sont généralement instantanées et suivent de très près l'émission de particules alpha ou bêta.

B. Les sources de radioactivité naturelle

Par la présence d'éléments radioactifs dans la croûte terrestre et l'existence de rayonnements cosmiques, l'homme vit dans un environnement naturellement radioactif. Ces deux sources de rayonnement naturel conduisent à une exposition interne et externe des individus. La composante dite « interne » de l'exposition correspond aux particules radioactives qui pénètrent dans l'organisme, elle constitue la part principale de l'exposition aux radiations ionisantes naturelles. Le radionucléide pénètre à l'intérieur de l'organisme le plus souvent par ingestion ou inhalation et continue à irradier au-delà du moment où a eu lieu l'incorporation. L'exposition est dite « externe » lorsque la source de rayonnements n'est pas en contact direct avec la personne et la dose reçue ne correspond qu'au temps pendant lequel s'est produite l'exposition.

En France, l'exposition annuelle aux rayonnements ionisants d'origine naturelle est de l'ordre de 2,4 mSv (Rannou et al., 2006). Elle correspond à la moyenne mondiale. Des régions présentent une exposition plus importante comme au Kerala en Inde (12,50 mSv) ou au Yanjiang, province du Sichuan en Chine (6,30 mSv) (NCRP, 2009; UNSCEAR, 2008). L'exposition annuelle moyenne aux rayonnements cosmiques a été estimée par l'UNSCEAR à 380 µSv. L'exposition aux rayonnements gamma telluriques a été estimée par l'UNSCEAR à 480 µSv.

B.1. Le radon

Le radon-222, gaz radioactif naturel, est produit par la désintégration du radium-226 descendant de l'uranium-238 présent dans les roches et les sols. Il est produit partout sur terre à partir de l'uranium contenu dans les formations géologiques constituant la croûte terrestre, en particulier celles qui sont proches de la surface. Une fois produit dans la roche, son état gazeux lui permet de circuler dans le sous-sol, de pénétrer et de s'accumuler dans les bâtiments. Il gagne rapidement la surface par les anfractuosités du sol, ou dilué dans l'eau des puits ou des nappes phréatiques et s'accumule dans les endroits peu aérés comme les caves et les sous-sols des maisons, etc. La nature des roches présentes sous les bâtiments est l'un des principaux paramètres déterminants du niveau de radon : les régions de France au sous-sol à forte teneur en uranium (Massif-Central, Corse, Bretagne) ont les concentrations les plus élevées. Les facteurs contrôlant sa concentration dans l'air intérieur sont multiples et dépendent en particulier des caractéristiques architecturales des constructions (matériaux de construction, existence d'un vide sanitaire…) mais également de la ventilation de celles-ci. La désintégration du radon provoque l'émission de particules alpha. En France, l'exposition annuelle au radon est de l'ordre de 1,43 mSv (Rannou et al., 2006).

B.2. Les rayonnements telluriques

L'écorce terrestre contient des atomes radioactifs qui sont la source d'une radioactivité naturelle : le thorium-232, l'uranium-235 et l'uranium-238. Dotés de très longues durées de

vie, de l'ordre du milliard ou de plusieurs milliards d'années, ils font partie intégrante de notre environnement. Avec leurs produits de filiation, ces radionucléides sont responsables du rayonnement tellurique qui varie beaucoup avec la nature du sol et certains types de roches. Celui-ci conduit surtout à une exposition externe au rayonnement gamma, aussi bien à l'intérieur qu'à l'extérieur des bâtiments, de l'ordre de 0,47 mSv par an en France (Rannou et al., 2006).

B.3. Les rayonnements cosmiques

Les rayonnements cosmiques proviennent de l'espace extra-terrestre et en particulier du soleil. Par interaction avec les noyaux d'atomes présents dans l'atmosphère, ils génèrent des rayonnements cosmiques dits secondaires (électrons, neutrons, protons, photons et mésons) dont l'intensité décroit jusqu'au niveau de la mer. Les rayonnements cosmiques possèdent une composante directement ionisante et une composante indirectement ionisante due aux neutrons, appelée composante neutronique, négligeable au niveau de la mer. L'exposition due aux rayonnements cosmiques est donc une exposition externe au rayonnement gamma, principalement. L'exposition annuelle moyenne est de l'ordre de 0,28 mSv en France (Rannou et al., 2006).

B.4. Les eaux minérales et les aliments

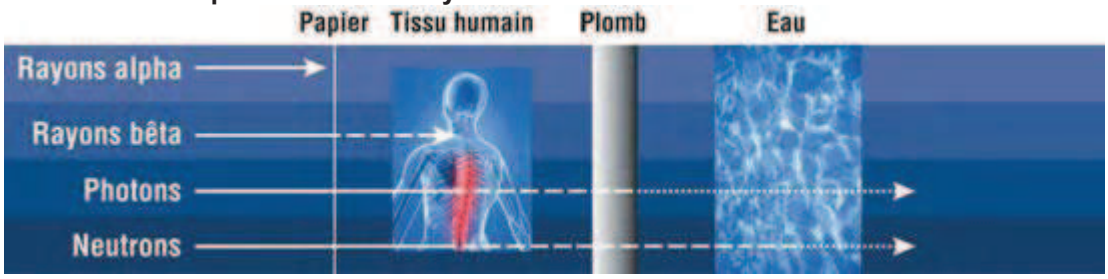
Les radionucléides de la croûte terrestre (potassium, principalement) et ceux créés par les rayonnements cosmiques (carbone 14, essentiellement) sont naturellement présents dans les plantes et les animaux, mais aussi dans l'eau. Nos aliments et nos boissons sont par conséquent légèrement radioactifs. Les radionucléides ingérés se fixent dans les tissus de l'organisme et entraînent une exposition interne de l'ordre de 0,23 mSv par an (Rannou et al., 2006).

C. Les doses et leur unité de mesure

L'activité est une caractéristique propre à un échantillon de matière contenant des noyaux radioactifs. Définie comme le nombre de désintégrations qui s'y produisent par

seconde, elle se mesure en Becquerel par L, kg ou m^3. L'activité d'une source ne prend pas en compte la nature et l'énergie des rayonnements alpha, bêta ou gamma. Par exemple, les particules alpha bien que rapidement arrêtées sont beaucoup plus dangereuses localement quand elles sont émises dans l'organisme que des rayonnements gamma dont l'effet sera réparti dans un plus grand volume (Figure 8).

Figure 8. Pouvoir de pénétration des rayonnements



Source :http://www.cnsc-ccsn.gc.ca/fra/resources/radiation/introduction-to-radiation/types-and-sources-of-radiation.cfm

Pour apprécier l'action de la radiation sur la matière, on utilise la dose absorbée. Il s'agit de la mesure de l'énergie que la radiation dépose par unité de masse de matière. L'unité de la dose absorbée est le Gray (Gy) qui correspond à une énergie déposée de 1 Joule (J) par kilogramme (kg). Le Gray ne tient pas compte directement du temps d'irradiation. Pour cela on trouve souvent les doses rapportées en Gray par unité de temps telles que le Gray/Heure et le Gray/An.

La dose équivalente est égale au produit de la dose absorbée moyenne dans l'organe, ou dans le tissu, par un facteur de pondération W_R qui tient compte de la nature du rayonnement (photons, électrons, neutrons, alpha) : facteur 20 pour les particules alpha très ionisantes localement, jusqu'à un facteur 1 pour les électrons, positrons et photons (X ou gamma) peu ionisants localement (ICRP, 2008). La dose équivalente s'exprime en Sievert (Sv).

La dose efficace, définie pour les besoins de la radioprotection, correspond à la somme des doses absorbées par tous les tissus, et s'exprime en Sv. Elle est doublement pondérée : une première fois par le facteur W_R et une deuxième fois par le facteur W_T permettant de tenir

compte de la radiosensibilité relative de chacun des tissus exposés. La dose efficace a été définie pour apprécier le risque total pour un individu résultant de l'exposition de l'ensemble de son organisme à un ensemble de sources de radiations ionisantes.

D. Exposition aux radiations d'origine naturelle en France

En France, comme dans d'autres pays et notamment en Europe (Gruber et al., 2013), l'exposition aux radiations d'origine naturelle a fait l'objet de plusieurs campagnes de mesures et études statistiques.

Concernant l'exposition au radon, un premier ensemble de données de mesures dans les habitations a été effectué au début des années 1980 et a inclus plus de 10 000 mesures de concentration de radon (Gambard et al., 2000). Deux études de cet ensemble de données ont été réalisées. La première consistait en une analyse statistique des caractéristiques des constructions modulant potentiellement l'activité volumique du radon, sur la base de 1 548 mesures réparties dans 15 départements (Valmari et al., 2011). La seconde a permis une estimation de l'exposition de la population française aux rayonnements naturels (Billon et al., 2005). Plusieurs études internationales ont montré l'influence des caractéristiques de l'habitat: le matériau de construction, le type de fondation, la période de construction principalement (Andersen et al., 2007; Borgoni et al., 2011; Gerken et al., 2000; Gunby et al., 1993; Hauri et al., 2012; Hunter et al., 2009; Kropat et al., 2014; Lehmann et al., 2002; Makelainen et al., 2001; Papaefthymiou et al., 2003; Pirard et al., 2000; Sundal et al., 2004; Zunic et al., 2007). Plusieurs études ont également montré la contribution de la géologie aux concentrations de radon à l'intérieur des habitations (Apte et al., 1999; Ielsch et al., 2002; Ielsch et al., 2001; Miles and Appleton, 2005; Minda et al., 2009; Sundal et al., 2004; Szabo et al., 2014; Zhu et al., 2001) avec comme principale source de radon le terrain sur lequel le logement est construit. En France, l'IRSN a mené plusieurs études sur l'influence des facteurs géologiques sur le taux d'exhalaison du radon et sur les concentrations de radon à l'intérieur de l'habitat (Ielsch et al., 2002; Ielsch et al., 2001; Mäkeläinen et al., 2009). Des recherches menées à l'IRSN ont permis de développer

une méthodologie de cartographie prédictive du potentiel d'exhalaison du radon à la surface des sols à l'échelle du territoire métropolitain (Ielsch et al., 2010). Ce type de cartographie a été développé dans d'autres pays comme en Allemagne (Kemski et al., 2001), en Belgique (Zhu et al., 2001), au Royaume-Uni (Miles and Appleton, 2005), en Irlande (Appleton et al., 2011), en Russie (Zhukovsky et al., 2011) et au Canada (Drolet et al., 2014). D'autres études, menées dans différents pays, ont utilisé des modèles statistiques pour évaluer l'exposition au radon à l'intérieur des habitations : modèles combinant différents paramètres tels que les caractéristiques du logement, du sol et du sous-sol, de la géologie (Andersen et al., 2007; Appleton and Miles, 2010; Friedmann and Groller, 2010; Hauri et al., 2012; Hunter et al., 2009; Kemski et al., 2009; Makelainen et al., 2001).

L'exposition aux rayonnements gamma telluriques et cosmiques, souvent étudiée conjointement à l'exposition au radon, moins variable à l'échelle d'un pays a fait l'objet d'un plus petit nombre d'études (Garcia-Talavera et al., 2013; Mora et al., 2007; Poje et al., 2012; Quindos Poncela et al., 2004; Rybach et al., 2002). En France, l'exposition aux rayonnements gamma a été étudiée par (Gambard et al., 2000; Rannou et al., 1985; Rommens et al., 2001).

Dans ce contexte, nous avons cherché à estimer l'exposition au radon et aux rayonnements gamma telluriques et cosmiques en France, à partir de données de mesures de concentration en radon et de rayonnements gamma fournies par l'IRSN ainsi que d'une cartographie prédictive du potentiel d'exhalaison du radon à la surface des sols.

II. Matériel

A. Données d'exposition au radon

A.1. Campagne nationale de mesures d'activité volumique (AV)

Une campagne nationale de mesures de l'activité volumique du radon (AV) a été menée en France avec pour premier objectif la connaissance de la distribution des activités

volumiques du radon (nombre de désintégrations par m^3, en Bq/m^3) dans l'habitat et l'estimation de l'exposition moyenne des populations des départements étudiés. Entre 1982 et 1990, 38 départements, choisis en fonction de caractéristiques géologiques favorables à l'émission du radon, ont fait l'objet de mesures. Entre 1992 et 1997, en collaboration avec les DDASS, l'Institut de Protection et de Sûreté Nucléaire (IPSN devenu depuis l'IRSN) et la Direction Générale de la Santé (DGS) ont relancé la campagne dans les 58 départements non mesurés dans l'objectif de couvrir l'ensemble du territoire. Enfin, en 1996, une nouvelle campagne de mesures a été relancée dans 22 départements avec un fort potentiel radon et peu de mesures. Avant 1991, la recherche des habitations où poser les dosimètres était laissée à l'appréciation des responsables de terrain de chacune des DDASS. Par la suite, la sélection des habitations dans un département s'est faite selon un critère géographique de répartition homogène des mesures et un critère démographique : un point de mesure pour 1 500 habitants (une mesure dans chaque commune de plus de 500 habitants en zone rurale ou entre 1 000 et 1 500 habitants en zone urbaine).

L'activité volumique du radon à l'intérieur des habitations était mesurée à l'aide d'un dosimètre Radon-Kodalpha qui se présente sous la forme d'un boîtier noir en plastique dont le couvercle sert de support à un petit film circulaire (LR115) capable d'enregistrer la concentration du radon présente dans la pièce où le dosimètre est posé. La durée de pose recommandée était de deux mois. Afin de préciser les conditions de pose du dosimètre et les caractéristiques de l'habitation où avait lieu la mesure, un questionnaire accompagnait chaque dosimètre. La base de données ainsi constituée devait aussi permettre d'améliorer la connaissance des facteurs contribuant à la variabilité des concentrations observées.

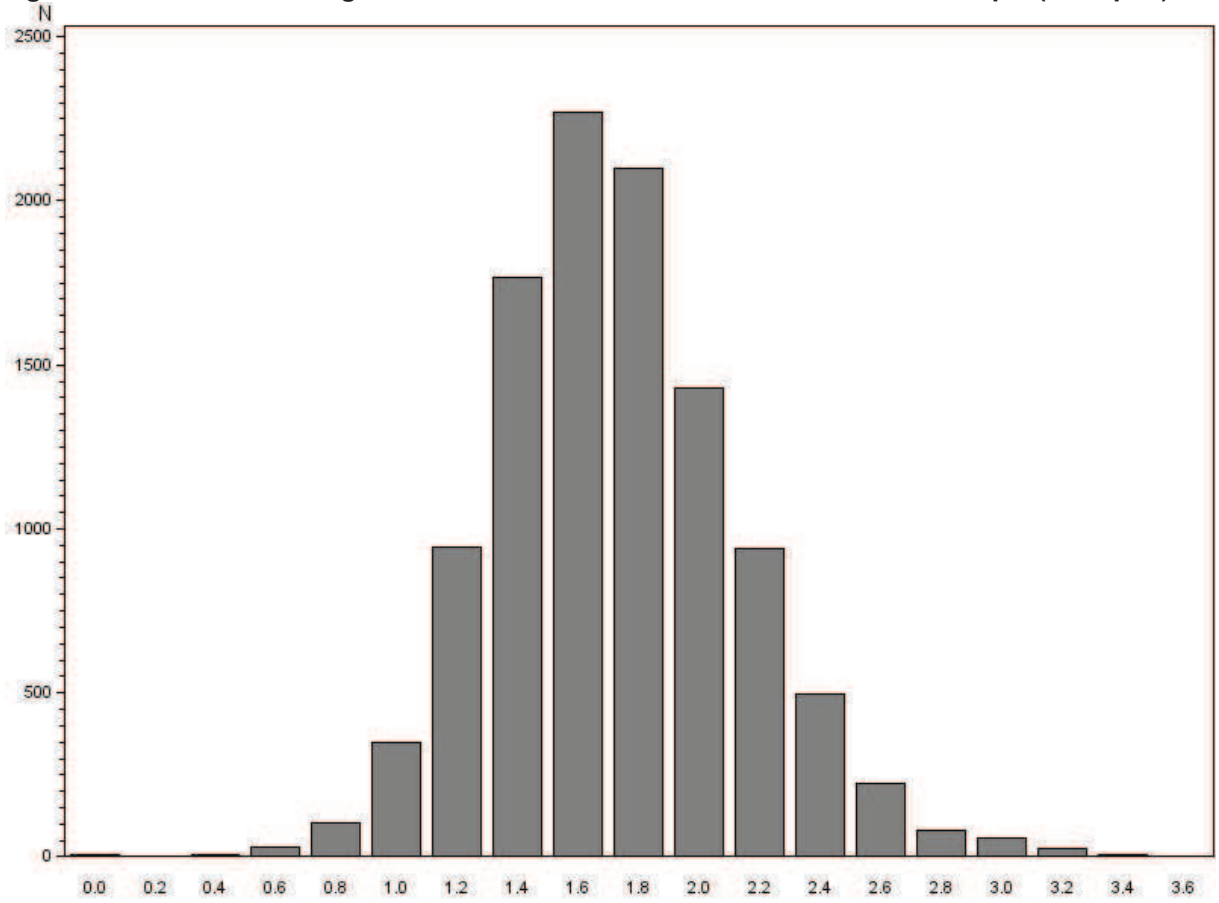
Les caractéristiques de l'habitat renseignées concernaient la période de construction (avant/après 1945/1948), le type de fondation du bâtiment (dalle, terre battue, vide sanitaire, sous sol, cave), le matériau de construction principal (bois, brique, béton, granit, parpaings, pierre), le type d'habitation (maison, immeuble, bâtiment public, école), le type de pièce (chambre, séjour, cuisine, cave), le niveau de la pièce (sous-sol, rez-de-chaussée, entresol, étage), le nombre de fenêtres et de portes vers l'extérieur, la présence de revêtement sur le

sol et le mur. Les conditions de pose du dosimètre étaient prises en compte par les dates de début et fin de la mesure, le nombre de jours de mesure, le degré d'aération (faible ou nul, moyen, élevé), le mode de ventilation (nulle, naturelle, mécanique collective ou individuelle), la présence de chauffage, de cheminée dans la pièce, la durée d'ouverture des portes et fenêtres.

Un ensemble de 15 759 mesures sont issues de la campagne. Parmi celles-ci, les mesures effectuées dans les caves ont été retirées car non représentatives de l'exposition résidentielle. De même, ont été retirées les mesures correspondant à une durée inférieure à deux mois ou supérieure à un an ainsi que les mesures redondantes (ex : confirmatoires). Un ensemble de 10 843 mesures d'activité volumique a ainsi été retenu pour l'analyse.

Les mesures d'AV dans les habitations ayant été réalisées à des périodes variables de l'année, une correction sur l'effet de la saison à laquelle a été effectuée la mesure a été appliquée à chaque mesure (Baysson et al., 2003), la saison pouvant influencer les mesures réalisées notamment du fait de variations saisonnières de l'aération des habitations. La distribution de forme log normale est donnée dans la Figure 9. La moyenne de cette distribution est 1,72 (moyenne géométrique 53,03 Bq/m^3), les 1er et 3ème quartiles respectivement 1,45 (28,29 Bq/m^3) et 1,97 (93,20 Bq/m^3).

Figure 9. Distribution du logarithme de base 10 des mesures d'activité volumique (en Bq/m^3)



Les abscisses 0 - 1,0 – 2,0 – 3,0 correspondent respectivement à des activités volumiques égales à 1 - 10 – 100 et 1 000 Bq/m^3.

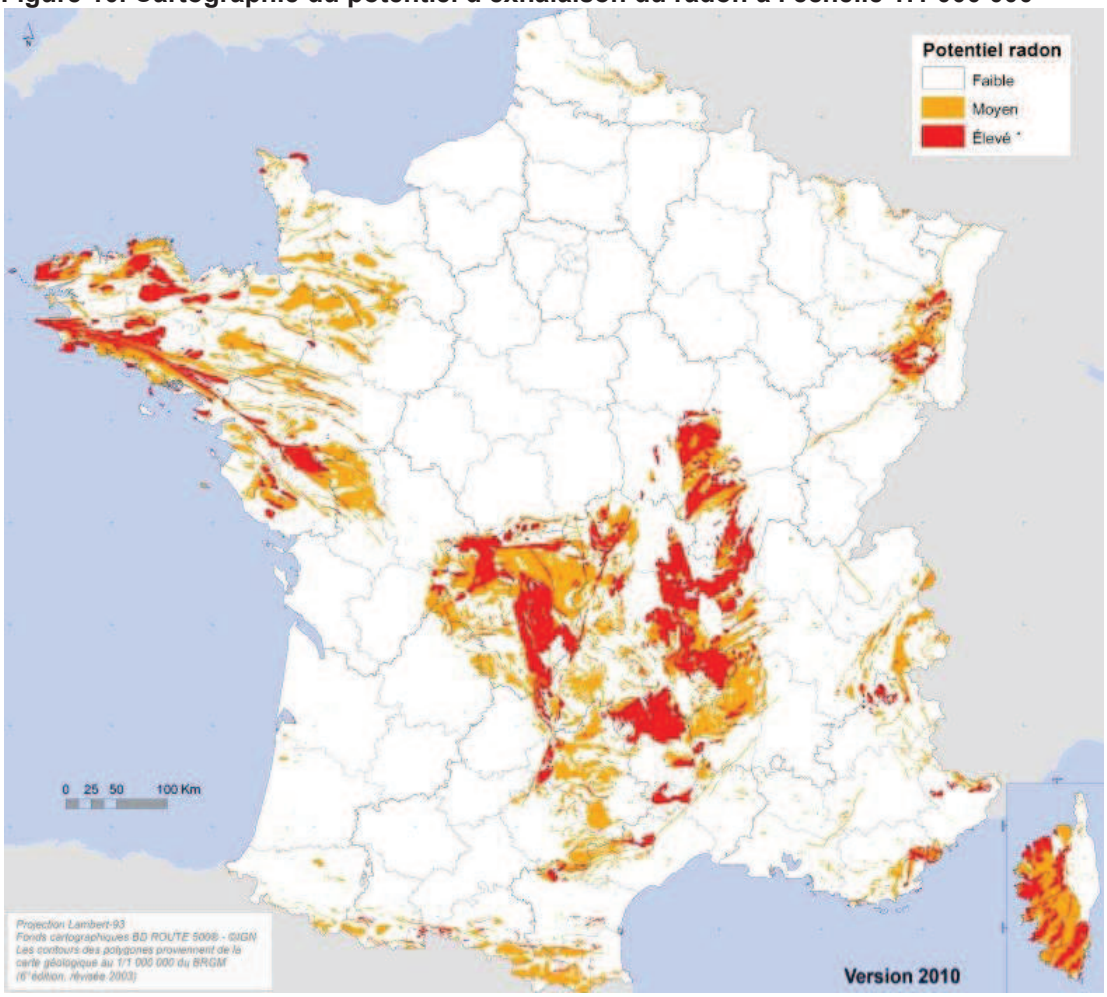
A.2. Potentiel d'exhalaison du radon (PR)

La connaissance des caractéristiques des formations géologiques sur le territoire rend possible l'établissement d'une cartographie des zones sur lesquelles la présence de radon à des concentrations élevées dans les bâtiments est la plus probable. C'est dans cet objectif que l'Autorité de Sûreté Nucléaire (ASN) a sollicité l'IRSN afin qu'il définisse et mette en œuvre une méthode permettant de cartographier le potentiel d'exhalaison de radon des terrains à l'échelle du territoire métropolitain.

Des recherches menées à l'IRSN ont permis de développer une méthodologie de cartographie prédictive du potentiel d'exhalaison du radon à la surface des sols. La démarche retenue vise à estimer le potentiel intrinsèque des terrains à générer du radon en surface en prenant en compte les principaux paramètres influençant la production du radon dans le sous-sol et le transport de ce gaz depuis sa source jusqu'à la surface des sols. Dans

une première étape, le potentiel du sous-sol en tant que source de radon est estimé et cartographié en attribuant une gamme de teneurs en uranium aux formations géologiques. Ce travail est réalisé à partir des sources d'informations disponibles: cartes géologiques, banques d'informations (analyses chimiques en uranium directement disponibles sur les formations géologiques de la zone étudiée), résultats de travaux de recherches (attribution d'une gamme probable de teneurs en uranium à une formation géologique présentant une lithologie donnée à partir des analyses de roches appartenant à cette classe lithologique). Il en résulte une catégorisation ordinale du potentiel source en radon en cinq classes. La seconde étape consiste à pondérer le potentiel d'émission de radon à la surface du sol en tenant compte de cofacteurs pouvant faciliter le transport du radon dans les roches et les sols : les failles majeures, les ouvrages miniers souterrains, les sites de sources hydrothermales. La classe de potentiel source dans la zone impactée étant alors incrémentée. Enfin, la dernière étape établit une classification finale de l'ensemble de ces paramètres pour aboutir à une carte du « potentiel radon » des formations géologiques à l'échelle 1:1 000 000. Il en résulte une catégorisation qualitative du potentiel radon sur l'ensemble du territoire français selon ces cinq classes: de faible à très élevé. Nous avons pu utiliser la cartographie du potentiel catégorisé en cinq classes mais celle-ci n'étant pas rendue publique, la Figure 10 présente une catégorisation en trois classes. La classe « faible » correspond au regroupement des deux classes d'exposition les plus faibles de la catégorisation en cinq classes et la classe « élevé » correspond au regroupement des deux classes d'exposition les plus élevées.

Figure 10. Cartographie du potentiel d'exhalaison du radon à l'échelle 1:1 000 000



source: IRSN/PRP-DGE/SEDRAN/BRN

B. Données d'exposition aux rayonnements gamma telluriques et cosmiques (RTC)

B.1. Campagne nationale de mesures

En France, des campagnes de mesure du débit de dose externe ont été réalisées à partir de 1977, en parallèle aux campagnes de mesure de concentration de radon. Le principal objectif était l'étude de la distribution des rayonnements à l'intérieur et à l'extérieur de l'habitat. Les mesures intègrent les rayonnements gamma d'origine tellurique et cosmique.

Les premières campagnes de mesure du débit de dose externe ont été effectuées de 1977 à 1990 ; trois types de dosimètres ont été employés (les tubes de sulfate de calcium, utilisés de 1977 à 1981 ; les dosimètres PGP1, de 1982 à 1985 ; les dosimètres

PANASONIC, de 1984 à 1990). Le principe de ces dosimètres largement utilisés en radioprotection réside dans leur propriété de restituer sous forme d'énergie lumineuse (émission de photons) lors d'un chauffage, l'énergie accumulée au cours de leur exposition par les interactions dues aux rayonnements ionisants. Durant les anciennes campagnes, les appareils de mesure ont été déposés à l'intérieur et/ou à l'extérieur des habitations, pendant une période d'environ six mois. En 1996, la DGS, les DDASS et l'IPSN ont relancé une nouvelle campagne de mesure du débit de dose externe dans l'habitat. Pour la nouvelle campagne, les mesures ont été réalisées avec un radiamètre SAPHYMO installé uniquement à l'intérieur des habitations, durant 45 minutes.

Dans certains départements, la dose enregistrée par les dosimètres ne prenait pas en compte les rayonnements cosmiques. Dans ce cas, la dose correspondante a été calculée grâce à la formule donnée par l'UNSCEAR 2000 et qui dépend de l'altitude.

$$D(z)=D(0)*(0{,}21*e^{-1{,}649*z}+0{,}79*e^{0{,}4528*z})$$

Où D(0) est la dose reçue au niveau de la mer (32 nSv/h) et z l'altitude en kilomètres.

Les données d'altitude ont été fournies par la société Géocible, l'altitude du point le plus proche sur une grille de 25 mètres de côté (BD ALTI) ayant été attribuée au point de mesure à partir de sa latitude et longitude obtenues par géolocalisation des adresses des habitations (cf Partie III. IV Géolocalisation).

L'ensemble final des données est constitué de 14 124 mesures, après suppression des mesures dont la durée d'enregistrement dépassait 1 an : 8 895 mesures réalisées à l'intérieur des habitations, 5 229 mesures réalisées à l'extérieur, réparties dans 60 départements.

B.2. Réseau de dosimétrie environnementale (dosimètres thermoluminescents TLD)

Le Laboratoire de surveillance atmosphérique et d'alerte (LS2A) de l'IRSN a pour mission la surveillance de l'environnement par dosimétrie passive. Un réseau de dosimétrie environnementale fut créé en 1985 en collaboration avec les DDASS et comprenait environ

1 200 sites en 2012. Le principe de couverture était principalement celui d'une surveillance environnementale de proximité des installations nucléaires pour contrôler l'éventuel impact de la radioactivité dans l'environnement. A partir de 2002, une recherche de nouveaux points a été entreprise pour une meilleure représentativité par région et département.

Les mesures étaient réalisées à l'aide de dosimètres thermoluminescents (Mechenet and Debayle, 2013). Les dosimètres étaient en moyenne renouvelés tous les six mois. L'ensemble des données est constitué de 16 628 mesures sur la période 2000 à 2012 réparties sur 1 191 sites géolocalisés à la mairie de la commune. Les mesures réalisées sur des sites non représentatifs de l'environnement, c'est-à-dire susceptibles d'être exposés à des niveaux de radiations élevés d'origine anthropique (aéroports, stations d'épuration, etc…) ont été retirées ainsi que les mesures dont la durée était inférieure à un mois ou supérieure à un an (du fait d'un potentiel problème de saturation du dosimètre). Les sites pour lesquels la variabilité des mesures était très importante par rapport à la variabilité intra site globale ont été étudiés et les mesures extrêmes considérées comme aberrantes, retirées. L'ensemble final est constitué de 12 699 mesures réparties sur 978 sites.

III. Méthodes

A. Estimation de l'exposition au radon

Deux estimateurs de l'exposition au radon ont été créés.

Le premier, le score d'exposition, est basé sur une exploitation de l'information apportée par la cartographie du potentiel radon à l'aide de disques formés autour des points de mesure de la campagne IRSN. L'intérêt de cette information est d'être disponible en tout point du territoire français, sa limite de ne prendre en compte que la composante géologique et en 5 niveaux seulement, des facteurs déterminant l'exposition au radon.

Le second estimateur est basé sur une estimation de l'exposition au radon faite par cokrigeage à partir des mesures de la campagne IRSN et de l'information supplémentaire apportée par le potentiel radon. L'intérêt de cet estimateur est d'être fondé sur plus de

10 000 mesures effectuées dans des domiciles et sa limite de nécessiter l'interpolation spatiale que propose le cokrigeage pour fournir une estimation en tout point du territoire.

A.1. Score d'exposition

A.1.a. Disques

Un score quantitatif S_k (PR ponctuel) a été affecté à chacune des cinq classes de PR ordinales (faible à élevé). S_k a été calculé comme la moyenne arithmétique des logarithmes décimaux des AV (logAV) mesurées dans des logements situés dans les zones où le PR a été classé dans la catégorie k. Un score a ainsi été attribué à chaque site où l'AV a été mesurée. Les valeurs moyennes des scores (PR moyen) ont été déterminées dans des disques centrés sur les points de mesure. Pour une habitation particulière i, le PR moyen dans un disque centré sur cette habitation a été calculé comme suit:

$$\text{PR moyen}_i = \frac{\sum_k A_{i,k} * S_k}{\sum_k A_{i,k}}$$

Où $A_{i,k}$ est la surface totale de PR de catégorie k dans le disque centré sur le logement i et S_k est le PR ponctuel dans la catégorie k.

Le PR moyen a été calculé autour des 10 843 points de mesures d'AV géolocalisés et pour des disques de rayons R = 100 m, 200 m, 500 m, 1 km, 2 km, 5 km, 10 km, 20 km, 50 km, 100 km, 200 km. Ces disques ont été analysés afin d'étudier lequel était le plus associé à l'AV.

A.1.b. Analyses statistiques

La distribution statistique observée des mesures d'AV étant très proche de la log-normalité (Figure 9, paragraphe II.A.1), les analyses statistiques ont été effectuées en utilisant les logarithmes décimaux d'AV.

Des modèles de régression multiple ont d'abord été utilisés afin d'évaluer l'effet des caractéristiques du logement et des habitudes de vie sur les mesures d'AV.

Des modèles de régression multiple ont ensuite été utilisés afin d'identifier la combinaison linéaire des buffers qui fournit le PR moyen le plus associé avec les AV. Le

modèle de régression multiple tient compte de l'existence d'autocorrélations spatiales des observations :

$$\log AV = \beta X + E$$

dans lequel:

- X est un vecteur de covariables pouvant inclure le PR, les caractéristiques de construction et le mode de vie des résidents,

- E est un vecteur d'espérance nulle et rend compte de la part de variabilité de logAV non expliquée par X. E inclut les erreurs de mesure ainsi que d'autres caractéristiques non prises en compte dans X, y compris certains aspects de la géologie, des caractéristiques du bâtiment et du mode de vie des résidents. Ces caractéristiques peuvent être très locales, si bien que lorsque la distance $d_{i,j}$ entre deux logements i et j tend vers zéro, la corrélation entre les valeurs de $\log AV_i$ et $\log AV_j$ ne tend pas nécessairement vers 1.

La corrélation spatiale entre $\log AV_i$ et $\log AV_j$ a été modélisée en utilisant une structure de corrélation exponentielle :

$$\text{corrélation}\left(\log AV_i, \log AV_j\right) = (1\text{-}n) * \exp\left(-\frac{d_{i,j}}{r}\right)$$

dans laquelle:

- $d_{i,j}$ est la distance entre les points de mesure i et j,

- n, appelé *effet pépite* en géostatistique, est la fraction de la variance spatiale liée à des facteurs locaux spatialement non structurés qui limite à une valeur inférieure à 1 la corrélation entre $\log AV_i$ et $\log AV_j$ lorsque la distance $d_{i,j}$ entre les points i et j tend vers zéro. Elle caractérise la variation entre deux mesures effectuées à des emplacements très proches.

- a, la *portée*, correspond à la distance à laquelle la corrélation entre deux mesures AV_i et AV_j est diminuée d'un facteur exp(-1) = 0,37.

Les régressions ont été effectuées avec la fonction GLS du package « nlme » du logiciel R[4] (version 2.13.1).

Pour tenir compte des corrélations possibles entre le PR et les caractéristiques du logement et du mode de vie des résidents, les moyennes d'AV entre les cinq classes de PR ont été étudiées pour chacune des différentes catégories de ces caractéristiques.

A.2. Cokrigeage

A.2.a. Interpolation

L'analyse descriptive de la variabilité spatiale et la possibilité de réaliser des interpolations spatiales font l'objet des principaux objectifs de la « Géostatistique ». Disposant de n valeurs mesurées Y_α en différents points α d'une variable Y « régionalisée » (c'est à dire *a priori* mesurable en tout point du territoire), il s'agit d'estimer au mieux la valeur prise par Y (notée Y*) en un point quelconque donné. Cette estimation est écrite comme une combinaison linéaire des valeurs observées Y_α. Dans cette combinaison linéaire le poids associé à chaque Y_α, noté ω_α, est calculé à partir d'estimations des corrélations spatiales entre Y* et les Y_α qui dépendent de leur position et de la structure générale des autocorrélations des Y (elle-même estimée à partir des mesures observées Y_α) (Wackernagel, 2003). L'équation correspondant à l'estimation de Y* est la suivante:

$$Y^* = \sum_{\alpha=1}^{n} \omega_\alpha Y_\alpha + \omega_0$$

Les poids ω_α vont donc incorporer les corrélations entre Y* et les Y_α, ainsi que celles entre les points Y_α et Y_β pour tous les couples (α,β). Ils sont choisis de façon à obtenir une prévision non biaisée et de variance minimale.

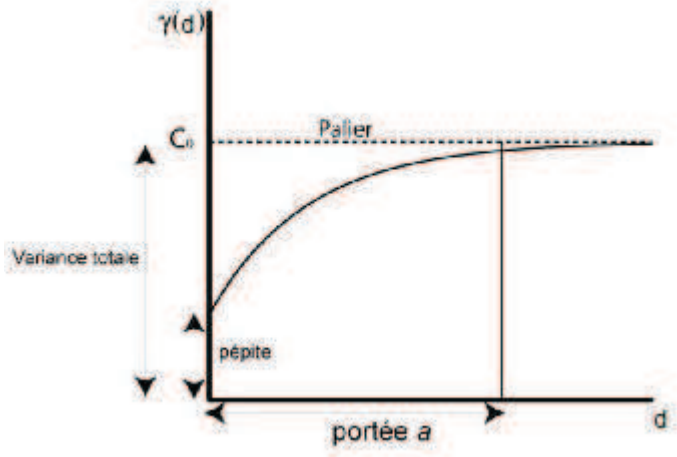
A.2.b. Semi-variogramme

Le semi-variogramme γ(d) est l'outil mathématique de base utilisé en géostatistique pour décrire les corrélations spatiales de la variable Y étudiée. Pour une distance d fixée,

[4] R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

γ(d) est égal à la moitié de l'écart quadratique moyen entre deux mesures en des points distants de d. La quantité théorique d'importance $\frac{E[Y(x+d)-Y(x)]^2}{2}$ est estimée en considérant tous les doublets (Y(x),Y(x+d)) distants de d et en prenant la moyenne des quantités $\frac{1}{2}$(Y(x),Y(x+d))^2 associées. Une fois cette opération faite pour différentes distances, est obtenu le semi-variogramme expérimental (Figure 11), auquel est ajusté un modèle paramétrique de la courbe γ(d). C'est sur la base du variogramme que peuvent être déterminés les coefficients à attribuer à chaque observation dans l'estimation de la variable d'intérêt en un nouveau point du territoire.

Figure 11. Semi-variogramme / Semi-variogramme expérimental



Abscisse : d = distance entre deux points. Ordonnée $\gamma(d)$ = moitié de l'écart quadratique moyen entre les Y_α et $Y_ß$ de tous les couples de points (α,ß) distants de d

Le semi-variogramme contient toutes les informations sur la structure de covariations spatiale des Y_α, supposée identique sur tout le territoire et ne dépendre que de la distance d entre deux mesures Y_α et Y_β considérées.

A.2.c. Krigeage simple

Le krigeage simple s'appuie sur la relation $Y^* = \sum_{\alpha=1}^{n} \omega_\alpha Y_\alpha + \omega_0$ à laquelle est ajoutée l'hypothèse que la moyenne de la variable Y est connue, et constante sur l'ensemble du territoire.

A.2.d. Cokrigeage

Le cokrigeage est utilisé lorsque deux variables sont disponibles : Y la variable que l'on cherche à estimer en tout point du territoire et T une covariable fournissant des informations sur Y. Dans l'estimation de Y, par exemple, l'équation devient :

$$Y^* = \omega_0 + \sum_{\alpha=1}^{n} \omega_\alpha Y_\alpha + \mu_0 \sum_{\beta=1}^{n} \mu_\beta T_\beta.$$

Cela nécessite de déterminer un modèle de variogramme pour Y et T, ainsi qu'un modèle de covariogramme, décrivant les corrélations entre ces deux variables.

A.2.e. Validation du modèle

Pour comparer les différents modèles de semi-variogramme entre eux, la méthode classiquement utilisée en géostatistique est la validation croisée. Elle consiste à retirer une valeur du jeu de données et à estimer la valeur retirée avec les toutes les autres données, puis à comparer la valeur estimée à la valeur réelle. La somme des carrés des écarts moyenne EQM calculée comme suit et rapportée à la variance totale donne une estimation de l'adéquation du modèle.

$$EQM = \frac{1}{n} \sum_{i=1}^{n} (Y_i - Y_i^*)^2$$

B. Estimation de l'exposition aux RTC

L'ensemble des mesures de RTC : 8 895 mesures de la campagne IRSN réalisées à l'intérieur de l'habitat, 5 229 mesures réalisées à l'extérieur de l'habitat et valeurs moyennes temporelles des mesures réalisées en 978 sites TLD a été étudié par unité géographique.

Le pourcentage de variance inter et intra unité a été calculé à l'aide d'un modèle à deux niveaux incluant l'effet de l'unité géographique j pour la mesure i:

$$Y_{ij} = \mu + \alpha_j + r_{ij} \text{ avec } \alpha_j \sim \text{iid Normale}(0, \theta_{00}) \text{ et } r_{ij} \sim \text{Normale}(0, \sigma^2)$$

θ_{00} représentant la variance inter unité et σ^2 la variance intra unité.

Tableau 8. Pourcentages de variabilité intra et inter unités géographiques expliqués en fonction de la finesse de la partition du territoire et du nombre de cas attendus [a] dans les unités disposant d'au moins une mesure.

Unité géographique	% variance inter	% variance intra	Nombre d'unités avec au moins une mesure (%)	Nombre de cas attendus (%) [a]
Zone d'emploi (N=304)	56,6%	43,4%	249 (81,9%)	8 184,5 (90,4%)
Bassin de vie (N=1 644)	58,9%	41,1%	1 209 (73,5%)	7 683,6 (84,8%)
Communes (N=36 570)	69,4%	30,6%	6 297 (17,2%)	4 172,4 (46,1%)

[a] Nombre de cas attendus (cf Partie III. II.B) et pourcentage du nombre total de cas attendus dans les unités géographiques disposant d'au moins une mesure.

Le pourcentage de variance intra unité était le plus faible pour la partition du territoire en communes mais celles-ci incluaient moins de 50% du nombre de cas attendus (Tableau 8). Les variances intra unités pour les *zones d'emploi* (ZE) et *bassins de vie* étaient assez proches. Par ailleurs, le nombre d'unités géographiques avec au moins une mesure ainsi que le nombre de cas attendus était plus important dans une description par ZE. Ce découpage a donc été retenu pour estimer l'exposition aux RTC à l'intérieur de l'habitat.

Dans les 167 ZE avec mesures intérieures et extérieures, la distribution conjointe des moyennes par ZE des mesures extérieures et intérieures réalisées montre une forte corrélation entre ces deux moyennes (Figure 12). La régression linéaire de la moyenne par ZE des mesures intérieures sur celle des mesures extérieures de la même ZE a permis une prédiction des moyennes par ZE des mesures intérieures.

Ainsi, dans les 55 ZE avec uniquement des mesures extérieures, l'estimation de l'exposition moyenne intérieure aux RTC a été estimée en utilisant un score linéaire de prédiction et la moyenne des mesures extérieures. Dans les 167 ZE avec mesures intérieures et extérieures l'exposition aux RTC a été estimée par une moyenne des mesures intérieures et de la prédiction faite par le score (avec une pondération par l'inverse de la variance des mesures dans la ZE). Enfin, dans les 27 ZE avec mesures intérieures et sans mesures extérieures : l'exposition aux RTC a été estimée par la moyenne des mesures de la ZE. Pour chaque ZE, la variance de l'estimation a été calculée à l'aide de la delta-méthode. Cette méthode propose d'approcher la fonction f(X) d'une variable aléatoire X par son développement limité à l'ordre 1 autour de sa valeur moyenne *m* :

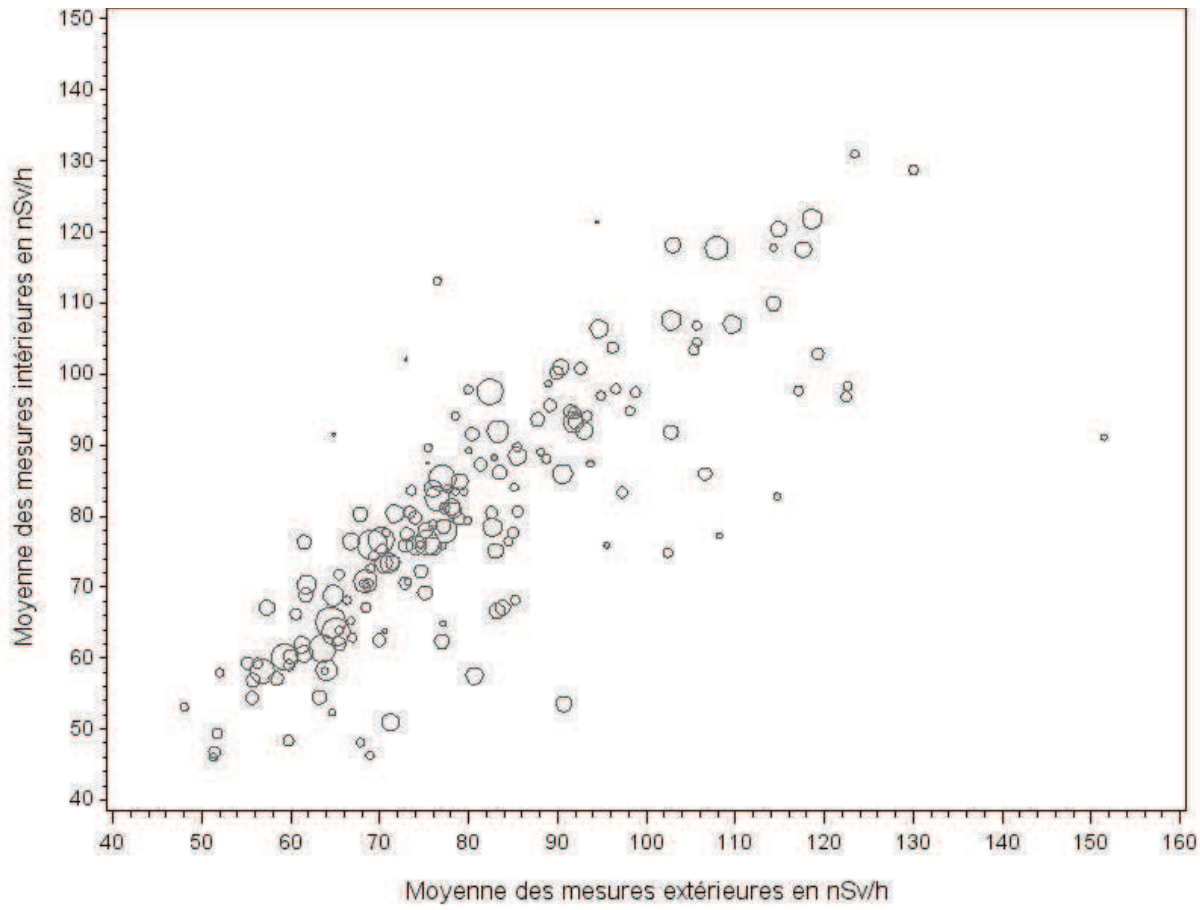
$$f(X) = f(m) + (X\text{-}m).f'(m) + 0(X).(X\text{-}m)$$

où O(X) est une fonction bornée de X ; ce qui permet d'écrire de façon approchée :

$$E[\,f(X)\,] \approx f(m) \text{ et } var[\,f(X)\,] \approx f'(m)^2 \,.\, var[\,X\,]$$

Approximation d'autant plus valide que la variance de X est faible, ce qui permet d'estimer le coefficient de variation de la moyenne des rayonnements gamma intérieurs par ZE.

Figure 12. Relation entre les moyennes des mesures intérieures et extérieures dans les 167 ZE disposant de mesures intérieures et extérieures



Moyennes géométriques par ZE. Si le nombre de mesures est 2 fois plus grand dans une ZE comparée à une autre, alors l'aire du cercle est 20 fois plus grande.

IV. Résultats

A. Exposition au radon

Afin de pouvoir mettre leurs résultats en relation avec les classes de PR correspondant aux lieux sur lesquels elles ont été effectuées, les mesures de la campagne nationale IRSN

ont été géolocalisées (cf Partie III. IV Géolocalisation). Des adresses postales suffisamment complètes pour mettre en œuvre cette démarche étaient disponibles pour 8 136 adresses. Les 2 707 mesures dont l'adresse comprenait uniquement la commune ont été géolocalisées à l'adresse de la mairie de la commune. Des analyses de sensibilité ont été réalisées afin d'étudier l'impact de la précision de géolocalisation :

- sur les 7 699 mesures géolocalisées à l'adresse : attribution du PR à l'adresse d'une part et à la mairie d'autre part.

- sur la totalité des mesures : attribution du PR à la mairie ou à l'adresse (si possibilité) d'une part et à la mairie pour la totalité d'autre part.

Les semi-variogrammes des deux ensembles présentaient les mêmes caractéristiques géostatistiques. L'échantillon des 10 843 mesures avec attribution du PR à l'adresse pour les mesures localisées à l'adresse et à la mairie pour les autres a donc été retenu.

A.1. Relation entre activité volumique, caractéristiques de l'habitat et potentiel radon

La répartition des mesures d'AV dans les classes de PR a été étudiée. Les moyennes d'AV des 10 843 mesures sont quasiment identiques dans les classes de potentiel 1 et 2 mais augmentent sensiblement de la classe 2 à 5. Le pourcentage de mesures ≥ 100 Bq/m^3 (resp. ≥ 400 Bq/m^3) croît de 18% à 69% (resp. de 2% à 31%) des classes de potentiel 1 à 5 (Tableau 9).

Tableau 9. Distribution des mesures d'AV par classe de potentiel radon

Classes	Nombre de mesures	Moyennes arithmétiques logAV (e-t)	Moyennes géométriques (Bq/m^3)	p20% - p80% (Bq/m^3)	% ≥ 100 Bq/m^3	% ≥ 400 Bq/m^3
1	5 352	1,67 (0,37)	46,9	23,8 - 93,2	18%	1%
2	3 887	1,68 (0,40)	48,1	23,1 - 97,2	19%	2%
3	1 010	1,91 (0,42)	82,0	36,1 - 176,5	40%	6%
4	450	2,13 (0,43)	134,2	56,8 - 288,7	60%	11%
5	144	2,27 (0,46)	185,5	72,8 - 463,6	69%	23%
Total	10 843	1,72 (0,41)	53,0	25,0 - 110,8	23%	2%

e-t : écart-type ; px% : x^{ème} percentile de la distribution

Nous avons également étudié la distribution des AV par catégorie de caractéristiques du logement et du style de vie. Les résultats sont présentés dans le Tableau 10. En ce qui concerne les matériaux de construction, la moyenne géométrique la plus élevée (171,5 Bq/m^3) a été observée pour la catégorie composée des 380 maisons construites en granit. Dans 18% de ces maisons, la concentration dépassait 400 Bq/m^3. Pour les autres matériaux, la moyenne variait de 36,5 Bq/m^3 pour le béton à 67,1 Bq/m^3 pour la pierre. En ce qui concerne le type de bâtiment, la moyenne était plus faible dans les 920 immeubles (30,9 Bq/m^3) comparée aux maisons. La concentration moyenne mesurée dans les bâtiments construits avant 1945/1948 (65,8 Bq/m^3) était supérieure à celle des logements plus récents. Pour les autres caractéristiques du logement (fondations, type de chambre et niveau du sol) les concentrations de radon moyennes étaient plus élevées pour les mesures effectuées au rez-de-chaussée (par rapport aux étages supérieurs), et pour les logements construits sur des dalles ou de la terre battue. En ce qui concerne la fréquence de renouvellement de l'air, les moyennes étaient relativement similaires pour les diverses catégories et la moyenne était, comme attendu, plus élevée quand il n'y avait pas de ventilation.

Tableau 10. Activité volumique de radon et caractéristiques de l'habitat

	Nombre de mesures	Moyennes géométriques (Bq/m^3) [IC95%]	≥ 100 Bq/m^3	≥ 400 Bq/m^3
Période de construction				
.	319	46,7 [42,7;51,3]	18%	1%
Après 1945/1948	5 669	44,4 [43,7;45,7]	17%	2%
Avant 1945/1948	4 855	65,8 [64,6;67,6]	31%	3%
Matériau de construction				
.	528	45,8 [42,7;49,0]	18%	2%
autre	289	56,7 [51,3;63,1]	25%	2%
bois	193	45,8 [39,8;52,5]	18%	1%
brique	2 489	40,8 [39,8;42,7]	14%	1%
béton	850	36,5 [34,7;38,9]	12%	1%
granit	380	171,5 [154,9;190,6]	68%	18%
parpaings	2 667	49,2 [47,9;51,3]	18%	2%
pierre	3 447	67,1 [64,6;69,2]	31%	3%
Fondations				
.	1 574	49,4 [46,8;51,3]	20%	2%
dalle, terre battue	3 677	62,4 [60,3;64,6]	29%	3%
vide sanitaire, sous sol	5 592	48,6 [47,9;50,1]	20%	2%
Type d'habitat				
.	167	45,0 [38,9;51,3]	17%	1%
autre	562	41,6 [38,9;44,7]	17%	1%
immeuble	920	30,9 [29,5;32,4]	8%	0%
maison	9 194	57,0 [56,2;57,5]	25%	3%
Type de pièce				
.	105	46,3 [38,0;56,2]	17%	3%
autre	1 066	46,5 [43,7;49,0]	19%	1%
chambre	1 113	50,4 [47,9;53,7]	21%	2%
séjour, cuisine	8 559	54,4 [53,7;55,0]	24%	2%
Niveau de la pièce				
.	221	50,8 [44,7;57,5]	23%	2%
rez-de-chaussée, entresol	8 394	56,1 [55,0;57,5]	24%	3%
étage	2 228	43,2 [41,7;44,7]	17%	2%
Fréquence d'aération				
.	402	41,2 [38,0;44,7]	14%	1%
abondamment	3 318	51,8 [50,1;53,7]	22%	3%
moyennement	6 731	54,6 [53,7;56,2]	24%	2%
peu, pas	392	51,7 [46,8;57,5]	22%	4%
Moyens de ventilation				
.	380	46,4 [41,7;51,3]	20%	3%
mécanique collective	212	33,8 [30,9;37,2]	7%	0%
mécanique individuelle	1 358	52,0 [50,1;55,0]	21%	2%
naturelle	3 646	48,7 [46,8;50,1]	21%	2%
néant	5 247	58,2 [56,2;60,3]	26%	3%
Total	10 843	53,0 [52,1;54,0]	23%	2%

A.2. Score d'exposition

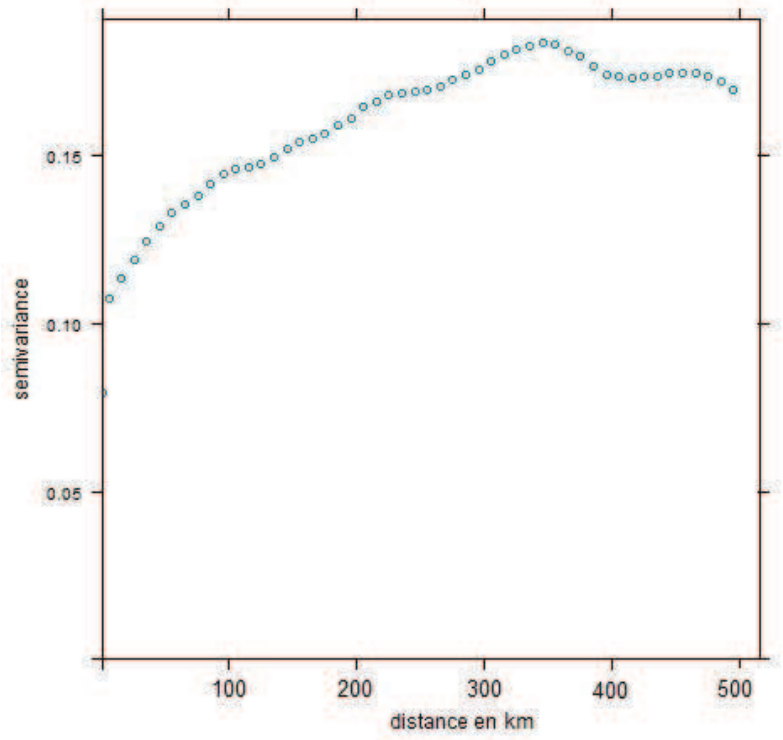
A.2.a. Activité volumique mesurée en un point et potentiel radon dans des disques centrés sur ce point.

Les corrélations entre les mesures d'AV en un point et les PR moyens déterminés dans des disques centrés sur ce point et de rayons 100 m à 20 km augmentaient de 0,31 à 0,40 puis décroissaient jusque 0,33 pour le disque de 200 km. L'information sur le PR contenue dans les disques de 20 km de rayon autour du point de mesure a été retenue. Pour tester la force de l'information à l'intérieur du disque, les modèles comprenant un disque de rayon $r < 20$ km et l'anneau compris entre ce disque et le disque de rayon 20 km ont été testés pour différentes valeurs de r par comparaison des rapports de vraisemblance, la combinaison de disques la plus corrélée avec l'AV était le disque de 2 km de rayon et l'anneau concentrique de 2 à 20 km.

A.2.b. Calcul du score d'exposition

Le semi-variogramme empirique des mesures d'activité volumique traduit l'existence d'une corrélation spatiale des mesures (Figure 13) puisqu'il croit progressivement avec la distance pour atteindre un plateau pour des distances élevées (la partie descendante du semi-variogramme observée au-delà d'environ 350 km correspond à des écarts quadratiques moyens calculés sur des effectifs de couple de points de plus en plus limités et donc à des estimations moins stables du semi-variogramme par le semi-variogramme empirique).

Figure 13. Semi-variogramme empirique des 10 843 mesures d'AV



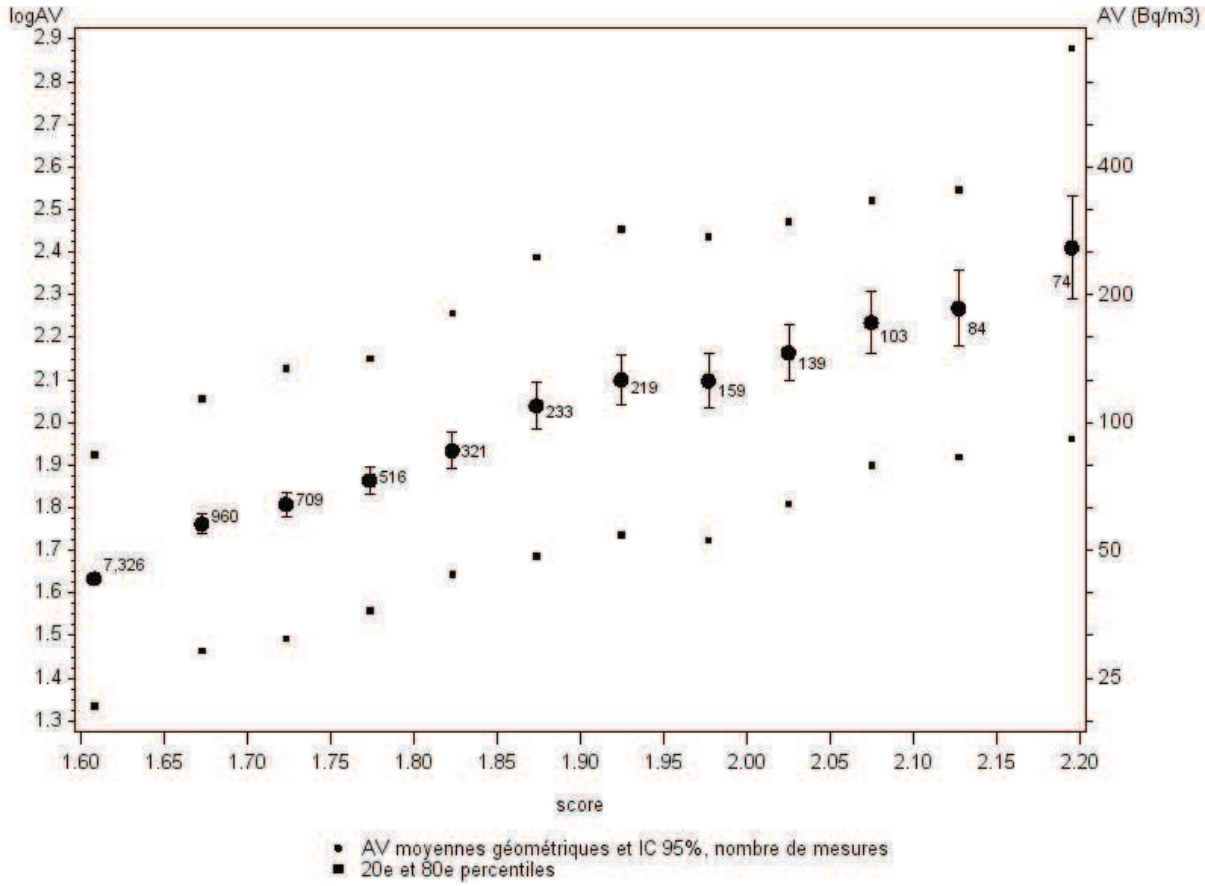
Cette corrélation spatiale des mesures a été prise en compte par un modèle de type exponentiel avec un effet pépite de 0,70 (la variabilité spatiale locale c'est à dire sans structuration spatiale exploitable pour estimer l'exposition au radon à partir des seules coordonnées spatiales a été estimée à 70%) et une portée de 82,7km.

Les coefficients de la régression des mesures d'AV sur le disque de 2 km et l'anneau allant jusque 20 km étaient 0,48 (IC95% [0,37 ; 0,59]) pour le disque de 2 km, 0,90 (IC95% [0,63 ; 1,16]) pour l'anneau et 0,70 (IC95% [1,10 ; 0,29]) pour l'intercept. Avec prise en compte de l'auto corrélation par un modèle exponentiel, 14,4% de variance était expliquée.

Un score a été calculé pour chaque point de mesure : Score d'exposition = -0,70 + 0,48 D_2 + 0,90 $A_{2\text{-}20}$ et découpé en 12 classes d'étendues égales (Figure 14). La moyenne géométrique de l'AV augmente de 43,1 Bq/m^3 (IC95% [42,3 ; 43,9]) pour la première classe à 257,8 Bq/m^3 (IC95% [195,1 ; 340,6]) pour la dernière classe. L'écart entre les valeurs des 20^e et 80^e percentiles est un peu plus important pour la dernière catégorie, mais stable dans les

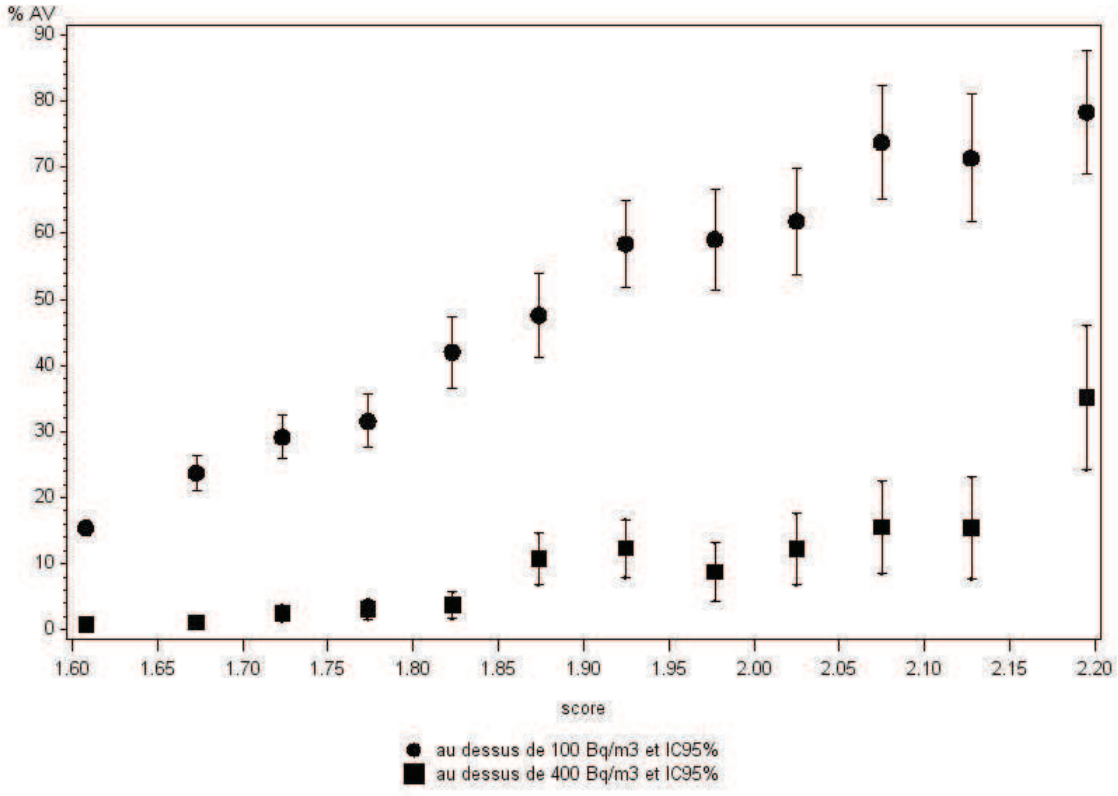
autres classes. Les AV varient d'un facteur d'environ 4 du 20^{e} au 80^{e} percentile dans chacune des classes.

Figure 14. Moyennes d'AV par classes de score de potentiel radon.



Le pourcentage d'AV au-dessus de 100 Bq/m^3 passe de 15,4% à 78,4% de la première à la dernière classe (Figure 15). Le pourcentage d'AV au-dessus de 400 Bq/m^3 passe de 0,8% à 35,1% de la première à la dernière classe.

Figure 15. Pourcentage de mesures d'activité volumique par classes de score de potentiel radon



Lorsque les caractéristiques de l'habitat sont incluses dans le modèle, 20,6% de la variance totale des mesures d'AV sont expliqués.

A.3. Cokrigeage

L'activité volumique du radon a été estimée par cokrigeage en tout point sur une grille de 5 km de côté (Pouchol and Ielsch, 2012). Le modèle de semi-variogramme retenu combine un modèle de forme exponentielle et un modèle sphérique. La valeur de l'effet pépite estimée est 0,56. Ceci indique comme pour le score d'exposition que la variance locale est élevée. L'erreur quadratique moyenne est 0,68 ce qui signifie que 32% de la variance totale des mesures d'AV sont expliqués par le cokrigeage des autres valeurs mesurées et du PR.

Les estimations cokrigées ont été obtenues aux coordonnées des 36 570 mairies françaises ainsi qu'aux domiciles des sujets GEOCAP. Ces estimations ont été obtenues avec leur erreur de prédiction qui varie notamment en fonction de la densité locale de points

de mesures. L'estimation de l'exposition au radon varie de 57,9 Bq/m^3 à 801,3 Bq/m^3 aux domiciles des témoins. La médiane de la distribution est 55,6 Bq/m^3. Les 1er et 3ème quartiles de la distribution respectivement 40,6 Bq/m^3 et 82,0 Bq/m^3 (Tableau 11).

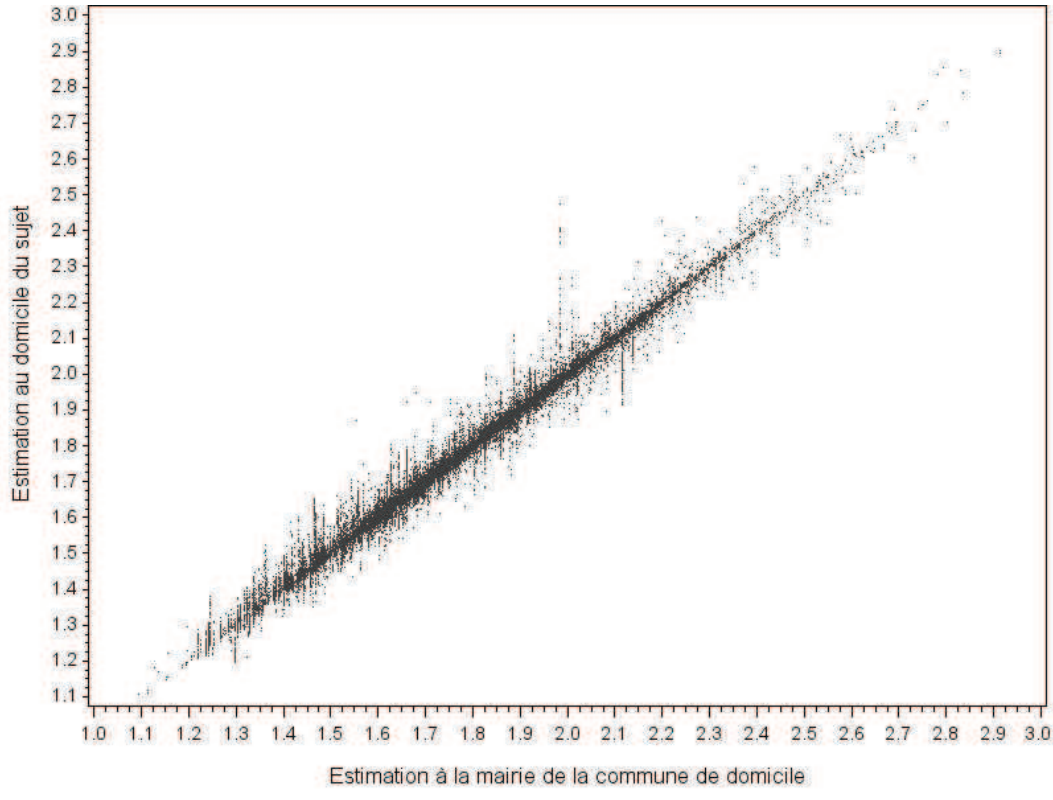
Tableau 11. Distribution de l'estimation de l'exposition au radon

Exposition en Bq/m^3	MG	e-t	min	p5%	p25%	p50%	p75%	p95%	max
Mairies des communes (N=36 326)	78,66	1,69	12,46	38,30	54,18	72,63	107,75	203,37	827,46
Témoins GEOCAP (N=30 000)									
domicile	57,92	1,72	12,76	24,92	40,63	55,60	82,02	145,32	801,25
mairie de la commune de domicile	57,34	1,73	12,46	24,30	40,37	55,26	81,59	145,62	819,16

MG : moyenne géométrique de la distribution; e-t : écart-type ; px% : x^{ème} percentile de la distribution

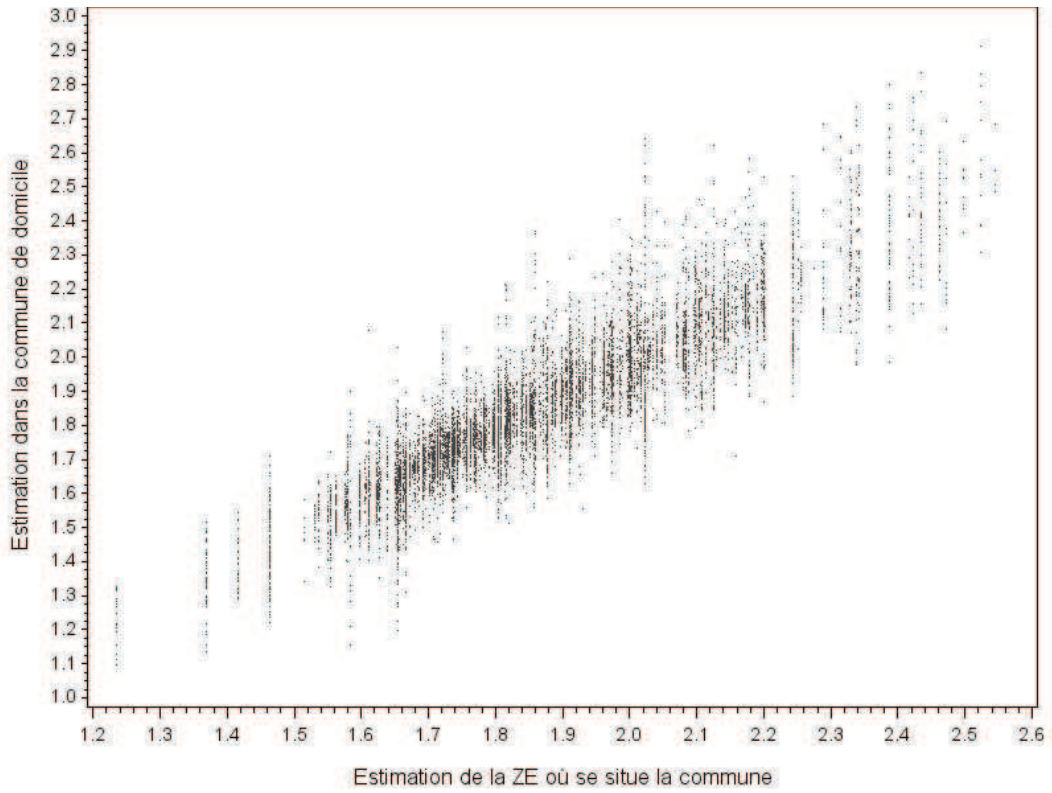
Comme le montre la Figure 16, les estimations cokrigées au domicile des témoins GEOCAP et celle de la mairie du domicile sont extrêmement proches avec une corrélation de Pearson égale à 0,99. Les estimations fournies par le cokrigeage ne peuvent donc pas être considérées comme estimant la concentration en radon au domicile mais plutôt comme une estimation de la valeur moyenne « lissée » de l'exposition au radon de cette commune.

Figure 16. Distribution conjointe du logarithme de base 10 des estimations des AV à la mairie et à l'adresse des domiciles des témoins GEOCAP



Comme le montre la Figure 17, les estimations cokrigées par commune sont, par contre, nettement variables entre les communes d'une même *zone d'emploi*.

Figure 17. Distribution conjointe du logarithme de base 10 des estimations des AV dans les ZE et dans les communes de résidence des témoins GEOCAP



B. Exposition aux RTC

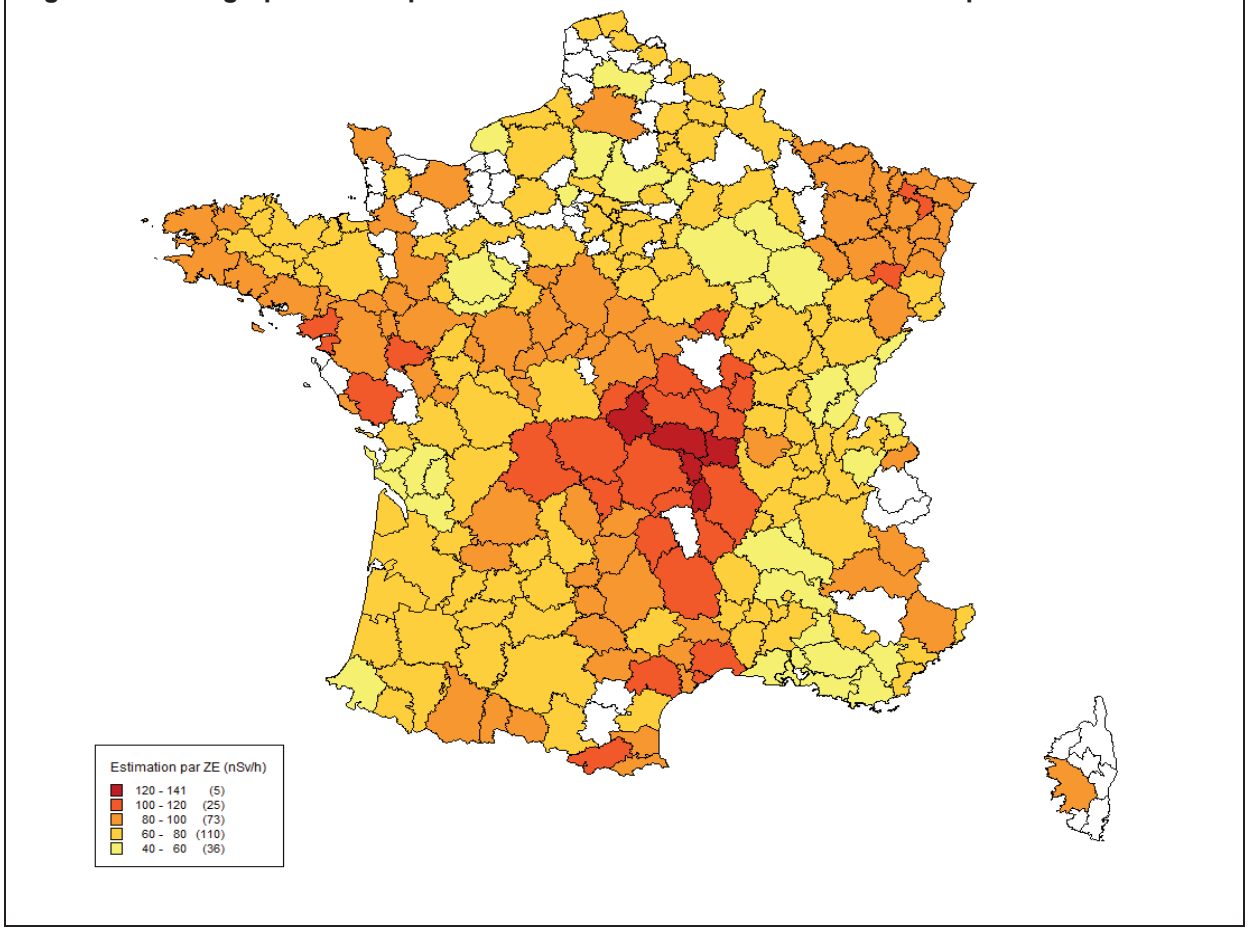
Dans les 249 ZE, l'exposition aux RTC varie de 46,21 nSv/h à 140,06 nSv/h dans le Massif Central (Tableau 12, Figure 18). L'exposition entre la ZE où l'exposition est la plus faible est celle où elle est la plus élevée varie d'un facteur 3 et d'un facteur 2 entre les percentiles 5% et 95% de cette distribution.

Tableau 12. Caractéristiques de l'exposition aux RTC à l'échelle des ZE

	MG	e-t	min	p5%	p25%	p50%	p75%	p95%	max
Exposition en nSv/h	78,18	17,85	46,21	52,01	65,02	76,19	89,75	110,92	140,06
Variance de l'estimation	2,69E-03	4,49E-03	1,73E-05	4,88E-05	1,12E-04	2,34E-04	1,46E-03	1,20E-02	1,81E-02
Coefficient de variation	1,90E-02	2,04E-02	2,16E-03	3,63E-03	5,61E-03	8,04E-03	2,03E-02	5,95E-02	7,14E-02

MG : moyenne géométrique de la distribution; e-t : écart-type ; px% : x^{ème} percentile de la distribution

Figure 18. Cartographie de l'exposition aux RTC à l'échelle des zones d'emploi



C. Exposition au radon et aux RTC

Afin d'étudier la corrélation entre expositions au radon et aux RTC, l'estimation de l'exposition au radon a été calculée pour chaque ZE par une moyenne de l'estimation de l'exposition au radon aux mairies pondérée par la population de la commune. Le Tableau 13 présente les résultats par ZE en fonction des classes d'exposition au radon. Ces deux expositions sont supposées liées. Toutefois, la corrélation entre les deux estimations est

égale à 0,43, il reste donc une part de variation de l'exposition aux RTC au sein des classes d'exposition au radon et notamment dans les classes de faible exposition au radon.

Tableau 13. Exposition aux RTC par classe d'exposition au radon

		Classes d'exposition au radon en Bq/m^3					
		<50	[50;100[	[100;150[	[150;200[	[200;250[	≥250
radon en Bq/m^3	nb ZE	67	158	55	12	8	4
	moyenne	40,35	69,25	121,81	174,35	222,36	309,32
	e-t	7,30	14,49	15,40	12,52	10,97	64,54
RTC en nSv/h	nb ZE	50	131	49	9	7	3
	moyenne	67,76	77,15	81,63	98,92	108,87	106,89
	e-t	12,39	14,31	19,00	19,89	28,79	12,26
	.	17	27	6	3	1	1
	<60	18	11	5	1	1	0
	[60;70[	15	31	9	0	0	0
	[70;80[	9	36	10	0	0	0
	[80;90[	4	27	10	0	0	0
	[90;100[	3	19	5	3	1	1
	[100;110[	1	5	6	2	2	1
	≥110	0	2	4	3	3	1

nb ZE : nombre de *zones d'emploi* ; e-t : écart-type

V. Discussion

L'exposition aux radiations d'origine naturelle a été estimée à l'aide de données fournies par l'IRSN pour l'ensemble du territoire français. Estimée à l'échelle des *zones d'emploi* dans une étude antérieure (Evrard et al., 2005), l'exposition au radon a ici été estimée à l'échelle des communes ou des domiciles. L'exposition aux RTC a pu être estimée à l'échelle des *zones d'emploi*, tandis qu'elle avait été estimée par département dans les études réalisées précédemment en France (Billon et al., 2005; Evrard et al., 2006a).

L'exposition aux sources artificielles d'irradiation n'a pas été considérée dans la présente étude. Elle provient des irradiations médicales diagnostiques (examens radiologiques, radiothérapie) (0,84 mSv par an), et des activités industrielles nucléaires et non nucléaires (0,03 mSv par an) (Rannou et al., 2006).

L'exposition aux radiations due à la partie neutronique des rayonnements cosmiques n'a pas été prise en compte dans la présente étude. La part neutronique des rayonnements cosmiques dépend principalement de la latitude. Les particules chargées de basse énergie sont, en effet, déviées par le champ magnétique terrestre et l'entrée des particules est

facilitée aux pôles. Elle est faible au niveau de la mer (Shahbazi-Gahrouei et al., 2013) et relativement invariante (UNSCEAR, 2006). Il est, en outre, relativement difficile d'estimer la part des rayonnements dus à cette composante (UNSCEAR 2000) et très peu d'études ont été réalisées (Schraube et al., 1998).

De la même façon, l'exposition aux radionucléides présents dans les eaux minérales et les aliments n'a pas été considérée mais celle-ci représente une faible part de l'exposition aux radiations naturelles.

L'exposition aux rayonnements gamma telluriques et cosmiques a fait l'objet de peu d'études : les caractéristiques de l'habitat ont un impact beaucoup plus fort sur les concentrations en radon dans l'habitat que sur les débits de dose des rayonnements gamma à l'intérieur de l'habitat et c'est principalement la localisation qui joue sur le niveau d'exposition aux RTC.

Le calcul de l'exposition au radon par le score d'exposition a ceci de spécifique qu'elle permet d'une part de tirer parti d'une information contextuelle relative aux potentiels prévalant autour des points de mesure mais également de tenir compte (dans une proportion toutefois difficile à quantifier) de la sensibilité des résultats aux erreurs de géolocalisation et aux imprécisions géographiques de la carte géologique au 1:1 000 000. Si des cartes géologiques (échelle 1:50 000 par exemple) et/ou des caractéristiques géologiques (par exemple des informations sur la perméabilité locale) plus détaillées avaient été utilisées, la géologie aurait probablement expliqué une plus grande proportion de la variation. Elle explique dans notre modèle de prédiction 14,4% de la variance. Lorsque sont ajoutés les caractéristiques de l'habitat, un pourcentage de 20,6% de variance expliquée est atteint. De façon semblable, d'autres pays, par exemple européens ont développé des modèles de prédiction. Par exemple, en Suisse (Hauri et al., 2013b; Hauri et al., 2012; Smith et al., 2007), 20% de la variance de la concentration en radon à l'intérieur de l'habitat a été expliquée par les données géologiques et les caractéristiques du logement, et en Angleterre (Hunter et al., 2009), 29 % de la variance de la concentration en radon a été expliquée par les unités géologiques et les paramètres du logement. Dans une autre étude, effectuée au

Danemark (Andersen et al., 2007), 40% de la variabilité de près de 3 000 mesures intérieures de radon a été expliquée par le modèle. C'est plus que dans la plupart des autres pays. Dans le modèle danois, les variables explicatives les plus importantes étaient le type de maison et la géologie. La géologie expliquait une partie de la variabilité similaire à celle expliquée par la géologie dans notre modèle alors que la variabilité expliquée par le logement était supérieure à celle expliquée dans la présente étude. Ceci peut être dû à une moindre capacité des caractéristiques du logement dont nous disposons de prendre en compte les variations à petite échelle dans les niveaux de radon ou à des caractéristiques du logement moins variables au Danemark. En Allemagne, des analyses de variance effectuées dans trois régions avec un nombre de mesures allant de 1 000 à 2 500 a montré que le sous-sol expliquait la part principale de la variabilité de radon à l'intérieur avec environ 36 % restés inexpliqués (Kemski et al., 2009).

Les résultats des méthodes géostatistiques génèrent des résultats complémentaires en ce qui concerne l'estimation de l'exposition au radon en France sous l'hypothèse que l'activité volumique moyenne est constante sur tout le territoire et ses variations modélisables par une structure d'autocorrélation homogène sur tout le territoire. La solution géostatistique renvoie à ce problème de non-stationnarité des valeurs de concentration de radon dans les habitations. La stationnarité correspond au principe selon lequel la corrélation entre les deux valeurs est déterminée par la seule distance géographique qui les sépare. Par exemple, deux maisons voisines partagent des environnements similaires (géologie, sol, des eaux souterraines, etc.) et devraient *a priori* avoir des mesures de concentration de radon similaires. Toutefois, les mesures peuvent aussi varier parce que les deux maisons ont des caractéristiques de construction très différentes ou leurs habitants des modes de vie différents, notamment du point de vue de l'aération du domicile. Il semble peu vraisemblable que la distribution spatiale de l'échantillon reflète parfaitement l'association entre deux caractéristiques (par exemple, la concentration de radon dans les habitations et le type de roche-mère). Ignorer la non-uniformité spatiale de l'échantillonnage des points de mesure et l'éventuelle non représentativité de ces emplacements peut entraîner des biais et donner

peu de sens à la comparaison entre les mesures de radon dans les habitations. L'un des intérêts majeurs de l'analyse géostatistique des variations spatiales de mesures est de prendre en compte cette variabilité.

Un peu plus de 32% de la variance sont expliqués par le cokrigeage. Une étude par région qui tiendrait compte des spécificités locales, si le nombre de mesures avait été plus important et avait permis une estimation robuste des semi-variogrammes empiriques régionaux, aurait probablement expliqué un plus grand pourcentage de variance (Shi et al., 2006). Dans le modèle de semi-variogramme retenu, l'effet pépite dont les origines physiques sont l'incertitude de mesure et l'existence de sources de variabilité locales vaut 56% de la variance totale. Cette valeur très forte montre combien le cokrigeage est contraint de « lisser ». Avec comme seules informations les coordonnées spatiales, même précises, c'est une moyenne lissée localement plutôt qu'une estimation ponctuelle locale qui est donnée. La présence d'un effet de pépite important n'est pas surprenante. Les semi-variogrammes des mesures de radon montrent habituellement un fort effet de pépite (Badr et al., 1996; Bertolo et al., 2009; Borgoni et al., 2010; Borgoni et al., 2011; Bossew et al., 2008; Dubois et al., 2007; Raspa et al., 2010; Zhu et al., 2001). En Belgique ou en Espagne, considérant un variogramme par type d'unité géologique, l'effet de pépite était présent bien qu'aucune corrélation entre mesures ne soit parfois mise en évidence (Cinelli et al., 2011; Garcia-Talavera et al., 2013). De plus faibles effets pépites ont été trouvés dans certaines études réalisées à une échelle plus petite et donc avec une plus petite variabilité (Buttafuoco et al., 2010; Franco-Marina et al., 2003). L'une des raisons d'un faible effet de pépite peut être le niveau de radon : pour des concentrations de radon de l'ordre de 100 Bq/m^3, l'incertitude de mesure peut être supposée petite.

La principale force de cette étude est sa couverture nationale, ce qui réduit considérablement le risque de biais de sélection. L'évaluation de l'exposition est fondée sur un modèle de prévision globale qui a été développé et validé à l'aide de plus de 10 000 mesures effectuées dans toute la France. Des études cas-témoins antérieures dans lesquelles les mesures étaient réalisées chez les sujets ont souvent rapporté une

participation inférieure à 55%, et les mesures d'exposition ont été souvent limitées à des sous-ensembles de participants à l'étude.

De l'analyse géostatistique, il ressort que la part de variabilité des mesures d'activité volumique radon qui est expliquée pas les coordonnées géographiques du point de mesure et les données sur le potentiel radon est relativement limitée. Le lissage effectué par le cokrigeage, permet de « capturer » efficacement cette part. Il ne permet pas, par contre, de fournir de l'information sur la part restante de la variabilité individuelle des expositions. Celle-ci n'est, en effet, explicable que par des informations relatives à chaque habitation et au mode de vie de ses habitants ; informations qui ne sont pas disponibles au niveau du registre, des recensements ou de l'étude cas-témoins GEOCAP.

Pour analyser l'association entre exposition aux radiations d'origine naturelle et leucémies de l'enfant, nous travaillerons à la fois sur une étude cas-témoins qui permet de prendre en compte les caractéristiques individuelles comme l'âge, la commune d'habitation et les éventuelles co-expositions. Nous travaillerons également sur une étude d'incidence à l'échelle des communes ou des zones d'emploi qui permet de couvrir un nombre d'années et de cas plus important.

Partie VI Radiations ionisantes d'origine naturelle et leucémies de l'enfant

I. Introduction

Les études épidémiologiques sur des groupes de personnes exposées à des doses fortes ou modérées de rayonnements ionisants ont montré que la leucémie infantile était particulièrement sensible aux irradiations (BEIRVII, 2006a; ICRP, 2008; UNSCEAR, 2008). Les radiations ionisantes, dont les utilisations, notamment médicales et industrielles, sont d'un grand intérêt pour l'homme, doivent donc faire l'objet d'une utilisation très étroitement maîtrisée, réglementée et contrôlée, afin de pouvoir apporter les bénéficies de leur utilisation sans être génératrice de risque pour l'homme. La « Radioprotection » vise à atteindre cet objectif et est fondée sur la combinaison d'informations dosimétriques, biologiques, cliniques et épidémiologiques.

Les études épidémiologiques des personnes exposées à ces rayonnements ionisants sont une source irremplaçable d'informations sur les risques de cancer liés aux radiations ionisantes chez les humains. La Life Span Study (LSS) est la cohorte des survivants japonais des bombardements atomiques d'Hiroshima et Nagasaki. C'est la source principale permettant d'estimer les risques cancérigènes liés à une dosimétrie précise des radiations reçues par une population humaine à des niveaux de risque directement observables et modélisables. De nombreuses données sur les personnes exposées pour des raisons médicales thérapeutiques ou diagnostiques offrent la possibilité d'étudier les risques causés par une exposition fractionnée à des doses thérapeutiques élevées. Les études de personnes exposées pour des raisons professionnelles et environnementales permettent une évaluation directe de l'exposition à faibles doses, et également de fournir des informations sur les différents types de rayonnement comme le radon et l'iode-131.

Le risque de leucémie infantile résultant de faibles doses de rayonnement reçues soit *in utero* soit dans les premières années de vie, du fait par exemple, d'une exposition médicale à visée diagnostique ou liée à des radionucléides présents dans l'environnement, suscite beaucoup d'intérêt (Wakeford et al., 2010). C'est d'autant plus le cas avec la fréquence croissante des méthodes modernes de radiographie médicale, comme la tomodensitométrie

pédiatrique (CT), qui a tendance à délivrer des doses plus élevées que les techniques radiographiques antérieures.

A. Cohorte des survivants japonais des bombardements atomiques d'Hiroshima et Nagasaki

La leucémie fut la première maladie maligne reconnue parmi les survivants japonais des bombardements atomiques d'Hiroshima et Nagasaki en Août 1945. C'est une des raisons pour lesquelles a été mise en place, à travers le recensement national japonais d'Octobre 1950, une cohorte de survivants japonais, la LSS. Une commission américano-japonaise est chargée de suivre cette population : la Fondation pour la Recherche sur les Effets des Rayonnements (Radiation Effects Research Foundation - RERF). Le suivi se poursuit encore aujourd'hui (Ozasa et al., 2012). La LSS comprend presque tous les survivants qui étaient proches des points d'impact (Preston et al., 1994) et se compose de plus de 93 500 survivants, dont environ 86 500 pour lesquels des estimations de doses de rayonnement reçues à la suite des bombardements ont été faites. Ces survivants ont reçu des doses délivrées brièvement, principalement de rayonnement gamma et de façon externe. Une gamme de doses étendue a été reçue par les survivants : 49 000 d'entre eux ont reçu des doses élevées de rayonnement (> 5 mSv) dont un peu plus de 2 000 personnes des doses supérieures à 1 Sv. Près des deux tiers des survivants exposés à moins de 5 mSv ont reçu des doses inférieures à 100 mSv (faibles doses). Entre 1950 et 1987, un total de 237 cas de leucémies a été rapporté parmi les membres de la LSS pour lesquels des doses avaient été estimées (Preston et al., 1994). Les données suggéraient que l'exposition aux radiations étaient à l'origine d'une grande part des leucémies apparues chez ceux qui ont été exposés à un jeune âge ; respectivement, 60%, 43% et 46% chez les personnes âgées de moins de 20 ans, entre 20 et 39 ans et de plus de 40 avec des distributions de doses similaires. 10 cas de leucémies de l'enfant ont été observés (Wakeford and Little, 2003) contre 1,6 attendu (Wakeford, 2013). L'excès de risque relatif est particulièrement prononcé pour un âge jeune à l'exposition : 6,1 par Sv pour les sujets exposés lorsqu'ils avaient moins de 20 ans

(UNSCEAR, 2000). L'analyse des observations indique une forte diminution de l'excès de risque avec l'âge à l'exposition ainsi qu'avec l'âge atteint (Richardson et al., 2009). En général, le pic de l'excès de risque est plus élevé et la décroissance du risque avec le temps est plus rapide pour ceux qui ont été exposés à un âge jeune (Preston et al., 1994).

Étant donné le RR élevé chez les enfants qui ont été irradiés après la naissance, une question qui reste sans réponse certaine est celle de l'absence de cas de leucémie infantile parmi les survivants qui ont été exposés *in utero* (Boice and Miller, 1999). Les survivants japonais irradiés in utero constituent une cohorte séparée de la LSS ; 800 d'entre eux ont reçu une dose de plus de 10 mGy *in utero* avec une dose moyenne de 250 mGy. Un faible nombre de cas (0,2 cas attendus) était donc attendu en l'absence d'exposition et il est possible que les cas incidents de leucémie infantile survenus entre 1945 et 1950, avant le début de la collecte systématique de données, n'aient pas été enregistrés.

Très récemment les données d'incidence ont été mises à jour pour la période 1950-2001 et incluent 312 cas de leucémies (LAL, LAM et CML) (Hsu et al., 2013). Des modèles avec excès de risque absolu (EAR) et excès de risque relatif (ERR) ont été considérés et les modèles ERR fournissaient des descriptions de qualité équivalentes aux EAR préférés dans les analyses antérieures. Les résultats montrent qu'une relation dose-réponse de forme linéaire quadratique s'ajustait mieux aux données qu'un modèle purement linéaire ou purement quadratique. L'excès de risque relatif était élevé pour l'incidence des leucémies 6,5 (IC95% [4,0 ; 10,3]), 3,9 (IC95% [2,3 ; 6,1]) et 4,0 (IC95% [2,1 ; 6,9]) par Gy chez les personnes âgées de moins de 20 ans, entre 20 et 39 ans et âgés de plus de 40 ans. Les analyses ont montré que le risque, en particulier pour les LAM, persistait tout au long de la période de suivi, 55 ans après l'explosion.

B. Les expositions médicales

La LSS ne peut pas fournir d'observations directes sur les risques dus à certaines modalités d'exposition ionisantes comme par exemple les expositions fractionnées de fortes doses de radiation (plutôt qu'une dose flash). L'observation d'autres populations est alors

essentielle (Gilbert, 2009). Les études portant sur des populations exposées à des rayonnements à des fins thérapeutiques ou à des fins de diagnostic (BEIRVII, 2006b; Ron, 2003; UNSCEAR, 2008) fournissent des observations directes sur ces modalités d'exposition.

B.1. Exposition à des fins thérapeutiques

La plupart des enquêtes épidémiologiques réalisées sur des groupes de patients ayant subi une radiothérapie, comme traitement d'une affection maligne ou bénigne (irradiation thymique, irradiation du cuir chevelu, traitement du cancer de l'enfant, irradiation pour hémangiome de la peau) ont confirmé le risque relatif élevé de leucémie infantile après irradiation dans les premières années de la vie qui ressort des données de la LSS (Wakeford, 2013). Ces études indiquent globalement que le risque de leucémie infantile augmente quelques années après l'irradiation, en soutenant l'hypothèse d'une courte période de latence déduite de l'expérience des membres de la LSS. La radiothérapie implique souvent un certain nombre d'expositions séparées dans le temps, de sorte que l'effet de ces doses fractionnées peut être comparé avec la seule dose reçue de façon brève dans la LSS.

Les résultats de ces études de radiothérapie vont dans le sens d'une augmentation du risque de leucémie (Doll and Wakeford, 1997; Hawkins et al., 1992; Kaldor, 1990; Ron et al., 1988; Shore et al., 2003; Tucker and Fraumeni, 1987) (cf Tableau 21 en annexe) mais présentent des limites et un certain nombre d'éléments doivent être pris en compte. Les radiations utilisées pour traiter la maladie, mais aussi la maladie elle-même, peuvent affecter les risques subséquents de cancer, de sorte que la généralisation à une population en bonne santé est incertaine. En outre, la radiothérapie nécessite des doses élevées conçues pour tuer les cellules anormales et ces doses sont souvent très localisées. De ce fait, les tissus à proximité des cellules cibles peuvent également recevoir des doses qui sont suffisantes pour tuer un nombre important de cellules normales. Cela conduit à une réduction du risque de cancer par dose unitaire dans ces tissus par rapport à celui résultant de doses faibles ou

modérées, comme une conséquence de la destruction de cellules. Ce peut être une question particulièrement importante concernant l'ampleur du risque leucémogène découlant de doses thérapeutiques reçues *in utero* ou peu après la naissance. Les doses spécifiques reçues dans les régions du corps éloignées de la cible de la radiothérapie (en grande partie due à la diffusion du rayonnement) sont des estimations de doses difficiles à calculer et font souvent défaut dans les études médicales ainsi que les coefficients de risque qui en résultent.

B.2. Exposition à des fins diagnostiques

Plusieurs études cas-témoins ont étudié l'incidence des leucémies en lien avec l'exposition aux radiations ionisantes utilisées à des fins diagnostiques. Dans une étude de cohorte de grande taille, l'Oxford Survey of Childhood Cancers (OSCC), Bithell and Stewart ont montré que le risque de leucémie augmentait de 47% pour des doses *in utero* de 10 mSv (Bithell and Stewart, 1975). En Nouvelle Zélande, 590 cas ont été diagnostiqués entre 1958 et 1961 et une association de l'incidence des leucémies avec les irradiations à visée thérapeutique mais pas à visée diagnostique a été rapportée. Bartley et al. et Infante-Rivrard et al. ont rapporté une association significative entre LA et exposition aux rayons X de diagnostic post-natal (Bartley et al., 2010; Infante-Rivard, 2003). L'augmentation de la fréquence d'utilisation des scanners CT qui délivre généralement des doses efficaces de quelques mSv, donne l'occasion d'évaluer directement les risques résultant d'expositions à faible dose. Plusieurs grandes études sont en cours et notamment en France avec une étude de cohorte rétrospective chez les patients de 14 grands services de radiologie pédiatrique (Bernier et al., 2012). Une évaluation de risque basée sur les pratiques médicales françaises a estimé que pour 100 000 scanners du crâne chez les moins de 5 ans, 4 cas de LA pouvaient être induits (intervalle d'incertitude à 90% [1 ; 14]) (Journy et al., 2014). Une étude de cohorte récente portant sur 178 000 jeunes âgés de 0 à 21 ans et ayant subit des scanners CT a montré une forte augmentation de risque mais non significative pour les leucémies (Pearce et al., 2012). Mathews et al. ont rapporté un risque de LA dans une

grande étude australienne basée sur 680 000 enfants ayant reçu une dose moyenne de 4,5 mSv par scanner (211 LA dans le groupe exposé, RR = 1,19 IC95% [1,03 ; 1,37]) (Mathews et al., 2013). Les études sur l'exposition post-natale aux rayonnements à visée diagnostique présentent encore des carences qui rendent l'interprétation des conclusions difficile. Les études ont utilisé des périodes différentes pour exclure les expositions qui se produisent au moment du diagnostic de la maladie, souffrent de la possibilité d'un biais de mémorisation car l'évaluation de l'exposition est fondée sur des données d'interview plutôt que sur les dossiers médicaux, et les différents types de procédures radiologiques peuvent conduire à des doses différentes (Wakeford, 2008). Dans l'ensemble, bien qu'il ne soit pas possible de tirer de conclusions définitives à partir de ces études, celles-ci offrent un certain support pour l'hypothèse d'un petit risque semblable à celui prédit par les modèles de risque classiques (voir, par exemple, les études récentes de (Rajaraman et al., 2011) et (Bartley et al., 2010)), mais ne fournissent pas de certitudes permettant de rejeter l'idée d'une absence de risque à faible dose (Wakeford, 2008).

C. Les expositions professionnelles

C.1. Expositions professionnelles des parents et leucémies chez les enfants

Des études ont été menées chez les enfants dont les mères étaient professionnellement exposées aux rayonnements ionisants pendant la grossesse (Bunch et al., 2009; Johnson et al., 2008). Ces études apportent des éléments de preuve, assez faibles, d'une possible influence des doses reçues *in utero* sur le risque de cancer chez l'enfant, un peu plus forte pour les cancers autres que la leucémie. Les doses intra-utérines moyennes de ces études étant très faibles (inférieures à 1 mGy), ces résultats ne doivent pas être sur-interprétés. Les résultats d'une étude sur les femmes qui ont été employées dans la centrale nucléaire de Mayak en Russie pendant leur grossesse ont été récemment publiés (Schonfeld et al., 2012). Parmi les 3 226 enfants exposés (ayant reçu une dose *in utero* moyenne de 54,5 mGy), 4 décès par cancer dont 2 de leucémie ont été identifiés chez les enfants, générant un ERR de 0,05 (IC95% [-0,0001 ; 1,334]) par mGy à la limite de la significativité statistique.

Bien que le petit nombre de décès sur lequel cette estimation est basée conduise à un intervalle de confiance étendu, la similitude de ce coefficient avec celui trouvé en utilisant des données de la OSCC est notable (Bithell and Stewart, 1975).

Dans l'installation nucléaire de Sellafield en Cumbria, en Angleterre, un lien entre l'exposition aux rayonnements ionisants des testicules de pères et le risque de leucémie chez leur progéniture a été suggéré pour la dose d'irradiation externe reçue avant la conception de leurs enfants (Gardner et al., 1990). Cependant, cette association, fondée sur seulement quatre cas de leucémie infantile (et aussi un petit nombre de témoins) avec une dose paternelle préconceptionnelle cumulée de plus de 100 mSv n'a pas été trouvée par les études épidémiologiques antérieures, notamment par les études réalisées chez les descendants des membres la LSS (Yoshimoto et al., 1990). Les résultats des principales études épidémiologiques sur l'exposition préconceptionnelle paternelle au rayonnement depuis l'étude de Gardner et al. ne soutiennent pas l'hypothèse d'un lien statistique substantiel entre LA et irradiation paternelle préconceptionnelle (Wakeford and Little, 2003).

C.2. Expositions professionnelles et leucémies de l'adulte

Les études concernant les expositions professionnelles chez les adultes sont considérées avec beaucoup d'intérêt dans la mesure où les conditions et les niveaux d'exposition ont été généralement recueillis avec précision. Ces études ont permis d'apporter des éléments complémentaires sur les effets des expositions externes chroniques à doses modérées. Les principales études portent sur les travailleurs de l'industrie nucléaire, les liquidateurs de Tchernobyl (personnel travaillant sur les opérations de gestion de crise post accident), les radiologues et techniciens en radiologie et les pilotes et personnels navigants de l'aérien. Les études sur le risque de leucémie chez les populations exposées professionnellement à la suite de contaminations accidentelles de l'environnement, en particulier les liquidateurs de Tchernobyl en Biélorussie, Russie et pays baltes (Ivanov et al., 2012; Kesminiene et al., 2008) et en Ukraine (Zablotska et al., 2013) signalent une augmentation du risque. Chez les travailleurs russes un ERR de 0,50 pour 100 mGy a été

signalé dans les 10 ans suivant l'accident (1986-1997) (Ivanov et al., 2012) et un ERR de 0,22 pour 100 mGy chez les travailleurs ukrainiens dans la période de 20 ans suivant l'accident (Zablotska et al., 2013). Les résultats des études de Tchernobyl portent sur des adultes ayant travaillé dans des environnements contaminés. Les études sur les populations de travailleurs nucléaires fournissent également des informations sur le risque dû à une faible dose. La plupart des estimations récentes (Muirhead et al., 2009) signalent une augmentation statistiquement significative du risque de leucémies autres que lymphoïdes chroniques. Une méta-analyse de 10 études sur les faibles doses reçues sur le long terme pour des expositions environnementales et professionnelles estime une relation dose-réponse statistiquement significative ERR=0,19 pour 100 mGy (IC95% [0,07 ; 0,32]) pour la leucémie, à l'exclusion des leucémies lymphoïdes chroniques (LLC) (Krestinina et al., 2013). Plusieurs études épidémiologiques sur les personnels navigants des compagnies aériennes ont été réalisées à travers le monde. Des estimations de la dose due aux rayonnements cosmiques ont été faites dans certaines études, basées sur un calcul annuel à partir des informations sur le nombre d'heures de vol, le type d'avion et les prévisions du vol. L'estimation de la dose annuelle varie entre 2 et 5 mSv et peut atteindre 75 mSv en fin de carrière (Zeeb et al., 2012). L'analyse d'une association entre les leucémies et l'exposition aux rayonnements cosmiques est disponible dans peu d'études (Gundestrup and Storm, 1999; Pukkala et al., 2003).

Pour des expositions modérées chroniques, une association significative entre la mortalité par leucémie et la dose externe de radiations a été observée dans la cohorte des travailleurs de Mayak (Shilnikova et al., 2003). L'ERR par Sv pour le risque de leucémie était comparable à celui observé dans la LSS.

Une analyse conjointe de 11 cohortes de mineurs de différents pays a été publiée en 1995 (Darby et al., 1995). En dépit de sa très grande taille, cette analyse ne comprenait que 69 décès par leucémie. Un risque significativement élevé de décès par leucémie est apparu, mais seulement pour les mineurs ayant travaillé moins de 10 ans. Une étude cas-cohorte rétrospective réalisée chez les mineurs d'uranium tchèques (basée sur 84 cas) conclut à une

association entre l'incidence des LA et l'exposition cumulée au radon (Rericha et al., 2006). Dans la cohorte tchèque des mineurs d'uranium (Tomasek et al., 2006), un excès significatif de mortalité par leucémie a été observé par rapport à la population générale (sur la base de 30 décès de leucémie). L'analyse a utilisé la dose équivalente cumulée à la moelle due au radon, à ses descendants, aux rayonnements gamma externes et à l'inhalation de poussières de minerai d'uranium. L'augmentation de la mortalité a été observée plusieurs décennies après l'exposition. Le risque accru de leucémie chez les mineurs d'uranium était significativement associé à la dose équivalente cumulée à la moelle, mais le radon y contribuait pour moins de 10%. Les résultats d'une étude réalisée sur la base des archives médicales des anciens mineurs d'uranium en Allemagne de l'Est incluant 377 cas de leucémie et 980 témoins ont suggéré un risque élevé de leucémie chez les employés avec une très longue carrière professionnelle dans l'exploitation minière d'uranium. Aucune association entre l'exposition à court terme aux descendants du radon et le risque de leucémie n'a été observée (Mohner et al., 2006). Une nouvelle analyse du même ensemble de données a été effectuée, en utilisant une méthode de calcul de la dose augmentée et en tenant compte des doses dues aux examens médicaux à rayons X (Mohner et al., 2010). Le pourcentage de la dose à la moelle absorbée résultant de l'inhalation de radon était d'environ 31 %. Un risque modérément élevé, mais pas statistiquement significatif, a été observé dans la catégorie des doses supérieures à 200 mGy (Mohner et al., 2010). Une analyse préliminaire du risque de leucémie a été réalisée dans le cadre du Projet européen Alpha-Risk. Au niveau mondial, un excès significatif de la mortalité par leucémie a été observé dans la cohorte français-tchèque-allemand (basé sur 69 décès par leucémie), en raison principalement de l'excès observé dans la cohorte tchèque (Marsh et al., 2008). Une association significative entre le risque de leucémie et la dose équivalente à la moelle a été observée, avec un ERR par Sv de 3,7 (IC95% [1,1 ; 8,8]). Certains éléments de preuve d'un risque accru de leucémie chez les cohortes de mineurs ont émergé dans les dernières années mais l'interprétation des résultats des études sur les mineurs est généralement compliquée en raison du faible nombre de cas et du risque de confusion entre les

expositions aux rayonnements gamma et l'inhalation de poussières minérales. Cette augmentation du risque semble être associée à une longue durée d'exposition, et une association avec la dose à la moelle est observée mais l'association n'est pas confirmée.

D. Les expositions environnementales

D.1. Expositions accidentelles

On dispose de peu d'observations directes sur le risque de leucémie après des expositions chroniques à des rayonnements ionisants à faible dose. L'étude sur les riverains de la rivière Techa en Russie contaminée depuis plusieurs décennies par les rejets radioactifs de l'industrie nucléaire en provenance du complexe militaro-industriel de Mayak, situé à proximité de la source de la rivière est à l'heure actuelle la plus informative pour la quantification du risque après une exposition environnementale (Krestinina et al., 2013). Des modèles d'excès de risque relatif ont estimé que près de la moitié des 72 cas de leucémie constatés entre 1953 et 2007 auprès des 28 223 membres de la cohorte étaient dus à l'exposition aux rayonnements ionisants. L'excès de risque relatif a d'abord été estimé à 0,22 pour 100 mGy. Cette étude montre que l'exposition à des radiations, pour des doses similaires à celles reçues par des populations exposées suite à des accidents nucléaires, est associée à une augmentation liée à la dose reçue à long terme du risque de leucémie. La mise à jour des estimations a mis en évidence un ERR de leucémie par unité de dose, moitié moindre de celles basées sur la dosimétrie précédente (Krestinina et al., 2010). Les études sur les populations exposées aux retombées de Tchernobyl n'ont pas montré d'augmentation de l'incidence des leucémies (UNSCEAR 2000, 2008). Les études publiées sur les adultes vivant dans la zone de Bryansk en Russie, Ukraine ou Biélorussie ont donné des résultats plutôt négatifs. Cardis et al. se sont intéressés au risque de LA en Europe et ont estimé que l'accident pouvait conduire à l'horizon 2065 à un excès global de 2 400 cas de leucémies soit une fraction attribuable de 0,04% (Cardis et al., 2006). Jusque 2005, l'estimation de la fraction attribuable est 0,08% avec un excès de 940 cas. L'UNSCEAR 2000, 2008 et le BEIR VII ont résumé les études sur l'incidence des leucémies chez les

enfants et adultes vivant dans les zones contaminées par l'accident de Tchernobyl et ont conclu que ces études ne supportaient pas l'hypothèse de risque de leucémie comme un résultat de l'exposition aux radiations ou, du moins, pas à un degré détectable.

D.2. Proximité des sites nucléaires

Concernant le rôle des radiations ionisantes à faible dose, les travaux les plus nombreux ont concerné le risque de leucémie et d'autres cancers à proximité des sites nucléaires. Dans la recherche des causes possibles de ces cancers, la plupart des études, qu'elles aient été locales et postérieures à la découverte de clusters ou multi-sites, ont le plus souvent abouti à des résultats négatifs mais ne permettant cependant pas d'écarter formellement l'hypothèse d'effet des radiations (Laurier and Bard, 1999). De nombreuses études épidémiologiques ont analysé l'incidence des leucémies de l'enfant autour de sites nucléaires en France (Bouges et al., 1999; Evrard et al., 2006b; White-Koning et al., 2004) et dans d'autres pays (Laurier et al., 2008) ou plus spécialement autour de certaines centrales nucléaires (Bithell et al., 2013; Heinavaara et al., 2010; Kaatsch et al., 1998; Spycher et al., 2011). Quelques excès nets et persistants ont été trouvés: autour des usines de retraitement britanniques de Sellafield (Bithell et al., 1994; Goldsmith, 1992) et Dounreay (Black et al., 1994) et autour de la centrale nucléaire allemande de Krümmel (Hoffmann et al., 2007; Kaatsch et al., 1998). Aucune hypothèse qui permette réellement d'expliquer ces associations n'a été identifiée si ce n'est un afflux de population à Sellafield et dans d'autres sites avec de très forts mouvements de populations. Par ailleurs, aucun excès de risque "général" au voisinage des CNPE (centrales nucléaires de production d'électricité) n'a été mis en évidence avant la publication allemande de 2008 sur les cancers de l'enfant à proximité des CNPE allemandes (Kaatsch et al., 2008; Spix et al., 2008). En France, deux précédentes études réalisées au sein de notre équipe ont analysé l'incidence des leucémies de l'enfant autour des sites nucléaires (Evrard et al., 2006b; White-Koning et al., 2004). Elles n'ont pas mis en évidence d'excès d'incidence des leucémies de l'enfant au cours des périodes 1990-1998 et 1990-2001, respectivement. Une nouvelle étude menée au sein de

l'équipe, motivée par la nécessité de tester en France les nouvelles observations Allemandes et portant sur 2 753 cas de LA pour la période 2002-2007 a utilisé la distance aux centrales nucléaires et un zonage géographique basé sur une estimation par modélisation de la dose à la moelle osseuse liée aux rejets gazeux des 19 centrales nucléaires françaises (Sermage-Faure et al., 2012). Un OR de 1,9 (IC95% [1,0 ; 3,3]), basé sur 14 cas, a été mis en évidence pour les enfants vivant à moins de 5 km des centrales nucléaires, par rapport à ceux qui vivent à 20 km ou plus loin, et une association très similaire a été observée dans l'étude d'incidence communale concomitante (SIR=1,9 IC95% [1,0 ; 3,2]). Néanmoins, en utilisant le zonage géographique fondé sur l'estimation de la dose à la moelle liée aux rejets gazeux, aucune association n'était retrouvée. Ces observations ont renforcé l'hypothèse selon laquelle l'excès d'incidence des leucémies de l'enfant observé sur 2002-2007 à moins de 5 km des CNPE n'était pas liés aux expositions aux radiations ionisantes liées à l'activité des CNPE. Elles conduisent au développement de recherches visant à identifier les facteurs, notamment démographiques, pouvant expliquer cet excès.

D.3. Expositions aux radiations ionisantes d'origine naturelle

Depuis la fin des années 1980, *une vingtaine d'études écologiques* ont mis en relation les variations géographiques d'incidence des leucémies infantiles et adultes avec les radiations d'origine naturelle. La plupart concluent à une corrélation positive (Laurier et al., 2001). Des revues portant spécifiquement sur les leucémies infantiles et la concentration de radon dans l'habitat dressent un constat similaire (Raaschou-Nielsen et al., 2008),(Tong et al., 2012). Un certain nombre d'études écologiques ont, en effet, étudié l'association entre les concentrations moyennes de radon et l'incidence de la leucémie standardisée sur l'âge (Alexander, 1990; Butland et al., 1990; Evrard et al., 2005; Evrard et al., 2006a; Foreman et al., 1994; Henshaw et al., 1990; Kohli et al., 2000; Lucie, 1989; Muirhead et al., 1991; Richardson et al., 1995; Thorne et al., 1996) ou du taux de mortalité (Collman et al., 1991) par zone géographique. L'exposition dans une région a généralement été mesurée comme

la concentration moyenne de radon dans l'air intérieur mais les concentrations dans l'eau de boisson (Collman et al., 1991) et au niveau du sol (Kohli et al., 2000) ont également été utilisées. La plupart des études portaient sur les leucémies pédiatriques dans leur ensemble mais les LAL et LAM ont également été étudiées séparément. Les associations écologiques ont été exprimées en coefficients de corrélation, coefficients de régression linéaire ou SIR. La plupart des études ont été menées au Royaume-Uni (Alexander, 1990; Foreman et al., 1994; Lucie, 1989; Muirhead et al., 1991; Richardson et al., 1995; Thorne et al., 1996) et d'autres en France (Evrard et al., 2005; Evrard et al., 2006a), en Suède (Kohli et al., 2000) et aux Etats-Unis (Collman et al., 1991). Parmi les 12 études, 11 ont montré une association positive entre le niveau de radon et les leucémies de l'enfant, 8 avec des associations statistiquement significatives ; une étude a montré une association inverse non-significative. Les quatre études portant sur les LAL ont montré des associations positives, trois d'entre elles statistiquement significatives. Les deux études portant sur les LAM ont montré des associations significatives.

Concernant l'exposition aux rayonnements gamma, l'étude de Henshaw et al. (1990) a montré une corrélation linéaire, positive et significative de 0,58 ($p<0,05$). Alexander et al. (1990), Lucie et al. (1990), Muirhead et al. (1991) et Richardson et al. (1995) ont observé une association écologique négative mais non significative entre l'incidence des leucémies de l'enfant et les rayonnements gamma dans les 22 comtés de Grande-Bretagne où l'exposition nationale moyenne (60 nGy/h ; UNSCEAR 2000) est plus faible que la moyenne mondiale (84 nGy/h).

En France, une étude écologique conduite pour la période 1990-1998 au niveau des *zones d'emploi* (Evrard et al., 2005) a montré une association modérée mais significative entre l'exposition au radon domestique et l'incidence des leucémies aiguës myéloblastiques chez les enfants : l'incidence était de 24% supérieure dans les *zones d'emploi* où l'exposition au radon était la plus forte (en moyenne, il y avait une différence de 100 Bq/m^3 entre le quintile d'exposition supérieur et le quintile inférieur pris comme référence). Aucune

association n'a été observée pour les rayonnements gamma dans l'étude conduite à l'échelle des départements pour la période 1990-2001 (Evrard et al., 2006a).

Les études cas-témoins réalisées en population générale ont rapporté des résultats contrastés et pour la plupart non significatifs (Raaschou-Nielsen et al., 2008). En 1987, Stjernfeldt et al. ont publié la première étude cas-témoins sur le radon mesuré à l'intérieur de l'habitat et la leucémie de l'enfant, avec 7 cas de LA vivant dans une région de la Suède (Stjernfeldt et al., 1987). La concentration moyenne de radon mesurée dans les dernières maisons occupées par les cas de leucémie était de 156 Bq/m^3 et 333 Bq/m^3 pour les témoins. Dans une étude réalisée aux Etats-Unis impliquant 942 cas et 1 292 témoins éligibles, les mesures de radon dans les maisons habitées pour plus de 70% de la période de l'enfance étaient disponibles pour 54% des cas et 34% des témoins. L'analyse donne un OR chez les LAL de 1,44 (IC95% [0,9 ; 2,3]) pour une exposition de 148 Bq/m^3 par rapport à une exposition de 37 Bq/m^3 (Lubin et al., 1998) et un OR de 1,02 (IC95% [0,5 ; 2,0]) sur la base de 281 cas (30 % des personnes éligibles) et 281 témoins appariés (22 % des personnes éligibles). En Allemagne, en Basse-Saxe, Kaletsch et al. ont étudié l'exposition au radon mesurée dans les résidences occupées pendant au moins 1 an chez 82 enfants atteints de LA (principalement des LAL) et 209 témoins correspondant à 40% de cas admissibles et 34 % des témoins admissibles. Les résultats ont montré un OR de 1,30 (IC95% [0,32 ; 5,33]) pour un contraste d'exposition supérieur à 70 Bq/m^3 par rapport à 70 Bq/m^3 et moins (Kaletsch et al., 1999). Steinbuch et al., ont étudié l'association entre le radon mesuré dans les maisons de 173 enfants diagnostiqués pour une LAM et 254 témoins au Canada et aux États-Unis (Steinbuch et al., 1999). Ces enfants correspondent à 80% des cas et 86 % des témoins admissibles pour les mesures et à 27% des cas et 33% des témoins disponibles pour l'interview. L'étude a montré un OR global de 1,1 (IC95% [0,6 ; 2,0]) pour une concentration de radon domestique de 100 Bq/m^3 par rapport à 37 Bq/m^3, et une association à la limite de la significativité après l'exclusion des enfants de moins de 2 ans. Les auteurs ont conclu que ce dernier résultat pouvait être dû au hasard. Au Caire, en Egypte, Maged et al. se sont intéressés au radon mesuré dans les

maisons occupées depuis la naissance pour 50 enfants atteints de LAL et 110 enfants témoins (Maged et al., 2000). L'étude a montré une association significative entre le radon et le risque de LAL avec des OR compris entre 5 et 7 pour la catégorie supérieure d'exposition. Ces estimations de risque sont de loin plus élevées que celles des autres études cas-témoins. Dans une vaste étude au Royaume-Uni, l'United Kingdom Childhood Cancer Study (UKCCS), le radon a été mesuré dans les maisons de 2 226 enfants atteints de cancer au moment du diagnostic (50% des cas admissibles) et 3 773 enfants témoins (31% des personnes éligibles) (UKCCS, 2002b). L'étude a montré une association négative avec les niveaux de radon pour les LAL avec un OR de 0,77 (IC95% [0,61 ; 0,99]) et pour les autres leucémies avec un OR de 0,71 (IC95% [0,43 ; 1,19]) pour les expositions supérieures à 30 Bq/m^3 comparées à celles inférieures à 8 Bq/m^3. Les auteurs de l'étude suggèrent que l'association inverse pourrait été liée aux différences de statut socioéconomique des ménages entre les cas et les témoins qui n'ont pas été observés en raison d'un taux élevé de non participation. L'UKCCS n'a pas mis en évidence d'association avec les rayonnements gamma pour les LAL (805 cas et 1 306 témoins) ou pour les autres types de leucémie (146 cas et 232 témoins) (UKCCS, 2002a).

Plus récemment, se sont développées des études fondées sur des enregistrements de cas et de témoins en population générale et utilisant des estimations d'exposition aux radiations ionisantes liées à la localisation des domiciles.

Une telle étude a d'abord été réalisée au Danemark pour tester l'hypothèse que l'exposition au radon domestique était associée à une augmentation du risque de leucémie et d'autres cancers infantiles (Raaschou-Nielsen et al., 2008). Les 1 153 cas incidents de leucémie diagnostiqués chez les enfants entre 1968 et 1994 ont été identifiés par le *Danish Cancer Registry*, et 6 697 enfants témoins ont été sélectionnés dans le *Danish Central Population Registry*. Un modèle de prévision a été utilisé pour calculer des estimations de concentration de radon dans les résidences des enfants depuis la naissance (Andersen et al., 2007). L'exposition cumulée au radon de chaque enfant a pu être estimée (somme des produits des durées d'exposition dans tous les domiciles occupés par les estimations des

niveaux d'exposition au radon dans ces domiciles). L'exposition cumulée au radon était associée au risque de LAL avec des RR de 1,21 (IC95% [0,98 ; 1,49]) pour une exposition comprise entre 0,26 et 0,89 kBq/m^3 - année et 1,63 (IC95% [1,05 ; 2,53]) pour une exposition supérieure à 0,89 kBq/m^3 - année par rapport à une exposition inférieure à 0,26 kBq/m^3 - année. Pour un enfant de 5 ans, cette différence dans l'exposition cumulée correspondrait à une étendue des concentrations en radon allant de 52 à 178 Bq/m^3. Considérant une dose-réponse linéaire, l'augmentation du taux d'incidence des LAL serait de 56% pour une augmentation de 1 kBq/m^3 - année. Aucune association statistiquement significative n'a été trouvée pour les autres types de leucémie.

Récemment, ont été publiés les résultats d'une étude nationale britannique cas-témoins sur l'exposition aux radiations d'origine naturelle basée sur les enregistrements des leucémies de l'enfant et d'autres cancers sur la période 1980-2006 (Kendall et al., 2012). L'étude a inclus 27 447 cas et 36 793 témoins appariés sur l'année et le registre de naissance, dont 9 058 cas de leucémie appariés à 11 912 témoins. L'exposition cumulée aux radiations était estimée à partir de l'adresse résidentielle de la mère à la naissance de l'enfant, en utilisant une carte prédictive basée sur des données de mesures de radon, et la moyenne des mesures prises par le *County District* pour l'exposition aux radiations gamma. Pour la leucémie infantile, un ERR de 0,12 par mSv de dose cumulée de rayonnement gamma a été observé (IC95% [0,03 ; 0,22]), alors que pour le radon, l'excès de risque était de 0,03 (IC95% [-0,04 ; 0,11]) par mSv.

En Suisse, une étude de cohorte (Hauri et al., 2013a) basée sur des données de recensement comprenant tous les enfants de moins de 16 ans à la date du recensement et comportant 997 cas de cancer dont 283 LA et parmi elles 225 LAL, n'a pas non plus observé de risque associé à l'exposition au radon. Basée sur une grille de résolution 2 km (Rybach et al., 2002), l'exposition aux rayonnements gamma a simplement été prise en compte comme une variable d'ajustement. Les niveaux de radon domestique ont été estimés pour chaque adresse en utilisant un modèle de prédiction sur la base d'environ 45 000 mesures effectuées dans toute la Suisse. Comparativement aux enfants exposés à une concentration

de radon inférieure à la médiane (77,7 Bq/m³), le RR ajusté pour les enfants avec une exposition supérieure au 90ème percentile (139,9 Bq/m³) était de 0,95 (IC95% [0,63 ; 1,43]) pour toutes les leucémies et 0,90 (IC95% [0,56 ; 1,43]) pour les LAL.

Au total, parmi les 6 études cas-témoins basées sur des mesures de radon réalisées dans les habitations, une seule étude a montré un risque significatif beaucoup plus élevé que les autres (Maged et al., 2000). Concernant les trois études fondées sur des enregistrements systématiques et sur un modèle de prédiction, l'étude danoise a rapporté une association significative avec les LAL ce qui n'est pas le cas des études anglaise et suisse. L'étude anglaise a quant à elle rapporté une association avec l'exposition aux rayonnements gamma.

Dans ce contexte, notre travail avait pour objectif de contribuer à la connaissance du risque de leucémie lié aux radiations d'origine naturelle en nous fondant sur deux études de grande taille réalisées à l'échelle de la France métropolitaine : une étude d'incidence sur la période 1990-2009 et une étude cas-témoins fondée sur des enregistrements exhaustifs de tous les cas de leucémie de l'enfant de 2002 à 2007 et d'un échantillon contemporain de 30 000 témoins représentatif des populations pédiatriques contemporaines.

II. Matériel et méthodes

L'échelle la plus fine à laquelle nous avons pu estimer l'exposition aux radiations ionisantes d'origine naturelle est celle de la commune, pour les expositions liées au radon, et de la zone d'emploi, pour les expositions liées aux radiations gamma telluriques et cosmiques. Au niveau géographique plus fin de l'étude cas-témoins GEOCAP, l'adresse des domiciles, la finesse spatiale des estimations élaborées est limitée par un inévitable lissage spatial (cf partie V. IV.A.). L'existence de facteurs de variabilité locaux non pris en compte par la seule adresse a donc réorienté notre démarche :

- pour bénéficier d'une puissance optimale nous avons développé une étude d'incidence et de corrélation écologique sur l'ensemble des 9 056 cas de la période 1990-2009,

- puis utilisé les informations géolocalisées à l'adresse de l'étude cas-témoins GEOCAP pour évaluer la stabilité de nos observations lorsqu'étaient exclus de l'analyse, les sujets proches de sources potentielles de risque de leucémies compte tenu de l'adresse précise de leur domicile.

A. Sujets

A.1. Etude d'incidence

L'étude d'incidence porte sur les 9 056 cas de LA, âgés de moins de 15 ans, enregistrés dans le RNHE pendant la période 1990-2009 parmi lesquels 7 434 LAL et 1 465 cas de LAM. L'âge considéré dans l'étude d'incidence est l'âge à la date de diagnostic en années révolues. A partir des données de recensement de l'INSEE et des interpolations que nous avons réalisées, les nombres de cas attendus de LA ont été calculés pour chacune des 20 années de la période par commune, sexe et année d'âge en utilisant comme taux de référence, les taux d'incidence de LA par sexe et année d'âge pour chaque année calendaire.

A.2. Etude cas-témoins

L'analyse cas-témoins porte sur la période 2002-2007 et inclut les 2 763 enfants atteints de leucémies aiguës diagnostiquées en France métropolitaine et qui avaient moins de 15 ans au 31 décembre de l'année du diagnostic. Les sujets témoins étaient les 30 000 enfants constituant l'échantillon témoin contemporain des cas, réalisé par l'INSEE, précédemment décrit et qui forment un échantillon représentatif de la population pédiatrique française (cf partie III. III) pour chacune des 6 années 2002 à 2007 (5 000 enfants par année).

B. Exposition aux radiations

L'exposition des sujets a été considérée de trois manières.

B.1. Exposition au diagnostic et exposition périnatale.

L'exposition au diagnostic correspond à l'exposition au radon estimée par le score d'exposition ou par cokrigeage dans la commune d'habitation au moment du diagnostic pour les cas, dans la commune d'habitation pour les cas attendus et témoins dans les études d'incidence et cas-témoins, l'unité est le Bq/m^3. Pour l'exposition aux RTC l'échelle d'estimation de l'exposition au diagnostic est celle des 304 *zones d'emploi* définies par l'INSEE en France métropolitaine (cf partie III. II) et l'unité est le nSv/h.

L'exposition périnatale pourrait avoir une importance potentielle sur les risques ultérieurs de leucémies de l'enfant (Wakeford, 2008). Nous ne disposons cependant pas de l'adresse du domicile des enfants à leur naissance dans les données utilisées pour l'étude d'incidence (RNHE et population INSEE) et pour l'étude cas-témoins GEOCAP. Cette information est néanmoins disponible pour l'ensemble des 764 cas de leucémies du RNHE diagnostiqués entre 2003 et 2004 et des 1 681 témoins de l'étude cas-témoins nationale Escale (Rudant et al., 2007a) réalisée en population générale. Les données ont été recueillies auprès des mères par une interview téléphonique standardisée. Nous avons étudié la mobilité résidentielle de cet ensemble de sujets (Tableau 14). Plus de la moitié d'entre eux n'a jamais déménagé (51%). L'estimation de l'exposition au radon étant réalisée au niveau communal, nous avons également considéré le pourcentage de sujets n'ayant jamais changé de commune et qui n'ont donc pas changé d'exposition dans notre étude. Chez les enfants âgés de moins de 4 ans, les plus nombreux, ce pourcentage varie de 89% chez les nouveaux nés à 60% chez les enfants de 4 ans. Au total, 59% des sujets n'ont pas changé de commune. Pour ces sujets, l'exposition cumulée correspond donc parfaitement à l'exposition au diagnostic multipliée par l'âge. Nous avons également étudié la corrélation entre l'exposition au radon estimée par cokrigeage pour la commune de résidence à la naissance et celle de la commune de résidence au diagnostic. Cette corrélation forte est assez stable selon l'âge du sujet et vaut 0,84 pour l'ensemble des sujets.

Tableau 14. Mobilité résidentielle et exposition au radon des sujets de l'étude cas-témoins Escale

Age au diagnostic (en années)	Nombre de sujets (cas/témoins)	Sujets n'ayant jamais changé de		Exposition au radon: corrélation entre la commune de résidence à la naissance et la commune de diagnostic [IC95%]
		logement	commune	
0	217 (30/187)	82%	89%	0,98 [0,97;0,98]
1	264 (82/182)	62%	75%	0,85 [0,81;0,88]
2	256 (103/153)	44%	64%	0,80 [0,75;0,84]
3	275 (109/166)	46%	63%	0,83 [0,79;0,86]
4	231 (86/145)	43%	60%	0,88 [0,85;0,91]
5	178 (62/116)	40%	62%	0,82 [0,77;0,87]
6	171 (59/112)	31%	54%	0,82 [0,77;0,87]
7	124 (42/82)	34%	52%	0,88 [0,83;0,92]
8	113 (32/81)	24%	45%	0,80 [0,72;0,86]
9	108 (33/75)	28%	50%	0,71 [0,60;0,80]
10	97 (31/66)	30%	57%	0,86 [0,80;0,91]
11	113 (29/84)	27%	43%	0,76 [0,67;0,83]
12	106 (25/81)	27%	47%	0,76 [0,66;0,83]
13	114 (23/91)	20%	48%	0,69 [0,58;0,78]
14	78 (18/60)	18%	36%	0,87 [0,79;0,92]
Total	2 445	51%	59%	0,84

Le même travail a été réalisé pour l'étude Estelle, enquête cas-témoins nationale qui inclut les enfants de moins de 15 ans atteints, entre 2010 et 2011, de leucémie, lymphome, tumeur cérébrale, neuroblastome, néphroblastome ou hépatoblastome. Les témoins ont été recrutés en population générale, selon une procédure qui assure la même répartition d'âge et de sexe que celle des cas. Les données ont été recueillies auprès des mères au cours d'une interview téléphonique d'environ une heure. Les résultats sont semblables : les corrélations entre l'exposition au diagnostic et l'exposition dans la commune de résidence à la naissance pour le radon, dans la *zone d'emploi* de résidence à la naissance pour les RTC sont très fortes (la corrélation entre l'exposition au radon estimée par cokrigeage pour la commune de résidence à la naissance et celle de la commune de résidence au diagnostic est 0,87). Les associations que nous aurions observées entre expositions aux radiations ionisantes au moment du diagnostic et leucémies de l'enfant sont probablement un bon reflet des associations que nous aurions pu observer en utilisant des informations sur les expositions à la naissance si nous en disposions.

B.2. Exposition cumulée

L'exposition cumulée d'un individu s'obtient comme la somme des produits de l'intensité d'exposition par la durée des expositions qu'il a rencontrées dans l'ensemble des domiciles qu'il a habités depuis sa naissance. Les expositions cumulées au radon et aux RTC sont respectivement exprimées en kBq/m^3 - année et en mSv. Elles dépendent de l'âge de l'individu et ont été estimées dans nos études en multipliant les valeurs de ces expositions au diagnostic par l'âge de l'individu. Bien que l'exposition cumulée reçue par un enfant dépende des différentes communes dans lesquelles il a résidé depuis sa naissance, nous ne disposions pas de ces histoires résidentielles pour les sujets de nos études. Toutefois, la corrélation observée sur les études Escale et Estelle entre les expositions à la commune de résidence à la naissance et à la commune de diagnostic pour les cas et à l'inclusion dans l'étude pour les témoins était bonne. Cela nous a conduits à estimer l'exposition cumulée aux radiations en multipliant exposition et âge au diagnostic.

B.3. Dose reçue à la moelle osseuse

Pour la leucémie, l'organe cible des radiations ionisantes est la moelle osseuse ; pour estimer la *dose reçue à la moelle osseuse*, il faut tenir compte du pouvoir de pénétration des rayonnements : les particules alpha sont moins pénétrantes que les rayonnements gamma. Pour tenir compte des dommages biologiques produits par ces rayons, il est également nécessaire de prendre en compte la nature du rayonnement: les particules alpha sont très ionisantes localement, et les électrons, positrons et photons (X ou gamma) le sont peu. Pour combiner convenablement ces deux composantes, des facteurs de pondération pour les rayonnements ont été appliqués. Pour estimer la dose à la moelle due au radon nous avons utilisé les facteurs de conversion concernant les concentrations de ce radionucléide dans les bâtiments (en Bq/m^3) à la dose à la moelle de l'enfant utilisés et publiés par Kendall et Smith (Kendall and Smith, 2002; Kendall and Smith, 2005). Les calculs des coefficients de dose (Sv/Bq) sont effectués à l'aide de modèles biocinétiques et dosimétriques qui sont utilisés pour décrire l'entrée dans l'organisme de radionucléides sous diverses formes chimiques

ainsi que leur distribution et leur rétention après leur entrée dans le sang. Des fantômes numériques, féminins et masculins, sont également utilisés pour calculer, pour toute une série de sources, la fraction d'énergie émise à partir d'une région source qui est absorbée dans la région cible. Ces approximations sont considérées comme étant acceptables pour les principaux besoins de la protection radiologique. Pour la dose à la moelle due aux RTC, nous avons appliqué des facteurs de conversion entre adultes et enfants pour la dose à la moelle : Petoussi et al. ont estimé que cette dose pour les enfants et les nourrissons était respectivement 12% et 24% plus élevée que celle des adultes (Petoussi et al., 1991). La dose totale à la moelle correspond au total de ces deux doses et s'exprime en mSv.

C. Analyses statistiques

Dans l'étude d'incidence, toutes les analyses statistiques concernant l'exposition au radon ont été réalisées à l'échelle des communes et celles concernant l'exposition aux RTC et à la dose à la moelle, au niveau des ZE. A chaque fois, les sujets et les unités géographiques ont été classés en catégories d'exposition de même étendue. Les analyses ont tout d'abord été réalisées pour tous les cas (0-14 ans) et pour la période entière 1990-2009. Elles ont ensuite été conduites séparément par classe d'âge (0-4, 5-9 et 10-14 ans), période (1990-1999, 2000-2009), et type de leucémie (LAL, LAM). Afin d'évaluer la stabilité des résultats et concernant l'exposition aux RTC, les analyses ont été réalisées en supprimant les *zones d'emploi* pour lesquelles l'estimation de l'exposition était la plus variable : les 63 *zones d'emploi* pour lesquelles le coefficient de variation de l'estimation moyenne aux RTC se trouvait dans le quartile supérieur de la distribution des coefficients de variations par *zone d'emploi* ont été exclues de l'analyse (cf Partie V. IV B). Cela correspondait aux unités géographiques dont le nombre de mesures était le moins important.

Des analyses ont également été effectuées en considérant une exposition *in utero* de 9 mois pour le calcul de la dose à la moelle.

Dans l'étude cas-témoins, les analyses statistiques ont été réalisées pour une exposition au radon faite au domicile du sujet et au niveau de la *zone d'emploi* pour

l'exposition aux RTC. Afin de tenir compte d'éventuels facteurs de confusion, des analyses supplémentaires ont été réalisées en excluant les sujets exposés à des expositions établies comme potentiellement liées au risque de leucémie dans l'étude GEOCAP.

D. Prise en compte des facteurs de confusion

D.1. Indice de défaveur au niveau contextuel

Un indice de défaveur a été utilisé afin de tenir compte du statut socio-économique des communes et des *zones d'emploi* pour l'étude d'incidence et des communes d'habitation des sujets de l'étude GEOCAP. Une analyse en composantes principales, pondérée par les PA de moins de 15 ans sur la période 1990-2009, a été réalisée au niveau des 36 326 communes de France métropolitaine sur les 4 variables suivantes : les pourcentages de chômeurs et d'ouvriers dans la population active des 15-64 ans, de bacheliers dans la population 1999 des plus de 15 ans et le revenu fiscal médian par Unité de Consommation indiqué par la Direction Générale des Impôts pour l'année 2001. 67% de l'inertie (variabilité) totale était restituée par la première composante principale.

Un indicateur de défaveur, basé sur les mêmes variables, a également été construit pour les années 1990 et 2006. L'indicateur basé sur l'année 1999 était corrélé à 0,77 avec l'indicateur de 1990 et à 0,83 avec l'indicateur de 2006. Pondérée par le nombre de PA de la commune, ces corrélations étaient respectivement de 0,93 et 0,95. Nous avons donc retenu l'indicateur de défaveur de l'année 1999, située au milieu de la période 1990-2009, pour ajuster les régressions de Poisson et régressions logistiques sur l'ensemble des périodes 1990-2009 et 2002-2007.

D.2. Co-expositions

Comme indiqué précédemment, afin de tenir compte d'éventuels facteurs de confusion, les expositions environnementales établies comme potentiellement liées au risque de leucémie dans l'étude GEOCAP ont été considérées. Concernant l'exposition aux champs magnétiques liés aux lignes à haute tension, un OR de 1,7 (IC95% [0,9 ; 3,6]) pour les sujets

GEOCAP résidant à moins de 50 m d'une ligne à très haute tension (225 kV ou 400 kV) comparés aux sujets vivant à plus de 600 m d'une commune où passait une telle ligne, a été observé (Sermage-Faure et al., 2013), en cohérence avec les différentes méta-analyses publiées sur ce sujet. Les 9 cas et 60 témoins concernés ont été exclus de l'analyse. Considérant une exposition plus large, une analyse a été effectuée en excluant les 176 cas et 1 840 témoins habitant à moins de 600 m d'une ligne à haute tension (63 kV à 400 kV). Concernant la proximité des centrales nucléaires, un OR de 1,9 (IC95% [1,0 ; 3,3]), basé sur 14 cas, a été mis en évidence pour les enfants vivant à moins de 5 km de l'une des 19 centrales nucléaires, par rapport à ceux qui vivent à plus de 20 km des 19 centrales nucléaires françaises (Sermage-Faure et al., 2012). Des analyses ont été réalisées en excluant les 14 cas et 80 témoins résidant dans un périmètre de 5 km autour d'une centrale nucléaire d'une part, les 137 cas et 1 061 témoins habitant à moins de 20 km d'autre part.

E. Puissance de l'étude

La puissance de l'étude à été testée en considérant l'hypothèse alternative que le risque de LA est celui déterminé par le modèle ERR décrit par l'UNSCEAR (UNSCEAR 2006) basé sur les données des survivants japonais de la LSS et calculé comme suit où *D* est la dose de radiation reçue en Sv, *a* l'âge atteint et *s* le sexe.

ERR : $h_0(s,a)*[1+(\alpha D+\beta D^2)*\exp(\kappa_1*\ln(a))]$

Les différentes hypothèses alternatives envisagées ont été simulées sous l'hypothèse H_1 suivante : $O_i \sim \text{Poisson}(\theta_i E_i)$ avec O_i et E_i les nombres de cas observés et attendus dans la l'unité i, θ_i est le risque calculé par le modèle ERR. Pour cela, le nombre total de cas O a été réparti selon une loi Multinomiale : $M\left(O, \frac{\theta_i E_i}{\sum \theta_i E_i}\right)$.

III. Résultats

A. Radon et leucémies de l'enfant

Le Tableau 15 montre les SIR obtenus dans l'étude d'incidence communale 1990-2009 pour toutes les leucémies aiguës, les LAL et les LAM âgés de 0 à 14 ans en fonction des

classes d'exposition au radon estimée au diagnostic. Aucune association n'est observée lorsque l'exposition au radon est estimée par cokrigeage à la mairie de la commune avec un découpage en classes de 50 Bq/m^3. Les SIR sont tous proches de l'unité et les résultats de la régression de Poisson n'indiquent ni hétérogénéité, ni écart à la linéarité, ni tendance linéaire significatives. Notons que pour les LAM, le SIR pour les 784 communes dont l'exposition est comprise entre 200 et 250 Bq/m^3 est significativement supérieur à 1 : 1,58 (IC95% [1,01 ; 2,34]) mais l'ajustement d'une régression de Poisson des LAM, linéaire par rapport à cette exposition, ne montre pas d'association avec un SIR pour 100 Bq/m^3 estimé à 1,02 (IC95% [0,91 ; 1,15]).

Lorsque l'exposition est déterminée par le score d'exposition découpé en classes d'exposition d'étendues égales sur le logarithme de base 10, aucune association entre les leucémies globales, les LAL ou les LAM n'est mise en évidence.

Aucune association n'a été observée pour toutes les LA par classes d'âge. Pour les LAM 0-4 ans, le SIR de la classe d'exposition 200 à 250 Bq/m^3 estimée par cokrigeage qui comprenait 16 cas observés et 7,1 attendus était significativement supérieur à 1 : SIR = 2,26 (IC95% [1,29 ; 3,68]) (Tableau 16). Cet excès significatif n'était pas retrouvé pour les âges supérieurs. Les résultats de la régression de Poisson sont à la limite de la significativité : exp(β)=1,14 (IC95% [0,97 ; 1,33]). Les analyses réalisées par sous période montrent un SIR de 1,94 (IC95% [1,08 ; 3,19]) pour 15 cas de LAM et 7,8 cas attendus âgés de 0 à 4 ans pour la période 1990-1999. Pour la période 2000-2009, celui-ci, basé sur 9 cas et 7,5 cas attendus n'est plus significatif (1,20 IC95% [0,55 ; 2,28]). Les SIR correspondant à la classe d'exposition de plus de 250 Bq/m^3 ne montrent aucun SIR allant dans le sens d'une association positive. Cette catégorie d'exposition ne regroupe néanmoins qu'un petit nombre de cas.

Afin de prendre ceci en considération, les communes appartenant aux catégories d'exposition « 200 à 250 Bq/m3 » et « plus de 250 Bq/m3 » ont été regroupées, comptant chez les 0-4 ans, 19 cas et 13,5 cas attendus de LAM pour la période 1990-2009. Il en résulte un SIR de 1,41 (IC95% [0.85 ; 2.20]). Les résultats de la régression de Poisson

étaient alors statistiquement significatifs : exp(β)=1,20 (IC95% [1,02 ; 1,40], p=0,03) et stables, bien que non significatifs, par sous-période (1990-1999 : exp(β)=1,20 (IC95% [0,96 ; 1,50], p=0,11 ; 2000-2009 : exp(β)=1,19 (IC95% [0,95 ; 1,49], p=0,13).

Au total, si la configuration des associations entre LAM et exposition au radon évoque la possibilité d'un lien positif, celui-ci manque de robustesse par rapport au découpage en classes d'exposition adopté.

Après ajustement sur l'indicateur de défaveur, ces observations restaient inchangées. Elles l'étaient également lorsqu'étaient exclus :

- les 9 cas et 60 témoins résidant à moins de 50 m d'une ligne à très haute tension (225 kV ou 400 kV) ou les 176 cas et 1 840 témoins habitant à moins de 600 m d'une ligne à haute tension (63 kV à 400 kV).

- les 14 cas et 80 témoins résidant dans un périmètre de 5 km autour d'une centrale nucléaire ou les 137 cas et 1 061 témoins habitant dans un périmètre de 20 km.

Tableau 15. Association entre l'exposition au radon et l'incidence de toutes les leucémies aiguës (N=9 056), des LAL (N=7 434) et des LAM (N=1 465) chez les enfants de 0 à 14 ans (France, 36 326 communes; 1990-2009)

Cokrigeage		Classes d'exposition en Bq/m^3						p_h [f]	exp(β) [g] IC95%
		<50	[50;100[	[100;150[	[150;200[	[200;250[	≥250		
	moyenne ± e-t [a]	40,6 ± 1,2	68,5 ± 1,2	120,8 ± 1,1	169,2 ± 1,1	222,4 ± 1,1	340,3 ± 1,3		
	moyenne pondérée [b]	34,9	67,5	120,4	168,1	221,2	332,8		
	N [c]	6 266	19 441	6 740	1 999	784	1 096		
LA	O/E [d]	3 732 / 3 788,7	3 957 / 3 877,2	941 / 974,2	253 / 235,3	100 / 94,5	73 / 86,2	0,20	1,00 [0,95;1,04]
	SIR [IC95%] [e]	0,99 [0,95;1,02]	1,02 [0,99;1,05]	0,97 [0,91;1,03]	1,08 [0,95;1,22]	1,06 [0,86;1,29]	0,85 [0,66;1,07]		
LAL	O/E [d]	3 049 / 3 108,2	3 279 / 3 184,3	762 / 799,9	209 / 193,2	75 / 77,6	60 / 70,8	0,12	0,99 [0,94;1,04]
	SIR [IC95%] [e]	0,98 [0,95;1,02]	1,03 [0,99;1,07]	0,95 [0,89;1,02]	1,08 [0,94;1,24]	0,97 [0,76;1,21]	0,85 [0,65;1,09]		
LAM	O/E [d]	624 / 614,8	606 / 625,7	161 / 157,4	39 / 38,0	24 / 15,2	11 / 13,9	0,26	1,02 [0,91;1,15]
	SIR [IC95%] [e]	1,01 [0,94;1,10]	0,97 [0,89;1,05]	1,02 [0,87;1,19]	1,03 [0,73;1,40]	1,58 [1,01;2,34]	0,79 [0,40;1,42]		

Score d'exposition		Classes d'exposition en Bq/m^3						p_h [f]	exp(β) [g] IC95%
		<50	[50;63[	[63;79[	[79;100[	[100;126[	≥126		
	moyenne ± e-t [a]	41,1 ± 1,1	55,4 ± 1,1	69,7 ± 1,1	88,5 ± 1,1	110,8 ± 1,1	143,1 ± 1,1		
	moyenne pondérée [b]	41,3	55,3	70,5	89,1	111,4	141		
	N [c]	29 859	3 168	1 551	871	582	295		
LA	O/E [d]	7 727 / 7 736	726 / 698	288 / 326,1	182 / 165,6	102 / 96,8	31 / 33,5	0,17	1,01 [0,86 ; 1,19]
	SIR [IC95%] [e]	1,00 [0,98;1,02]	1,04 [0,97;1,12]	0,88 [0,78;0,99]	1,10 [0,95;1,27]	1,05 [0,86;1,28]	0,93 [0,63;1,31]		
LAL	O/E [d]	6 361 / 6 350,1	585 / 573,1	223 / 267,8	154 / 136	87 / 79,5	24 / 27,5	0,05	1,01 [0,83 ; 1,19]
	SIR [IC95%] [e]	1,00 [0,98;1,03]	1,02 [0,94;1,11]	0,83 [0,73;0,95]	1,13 [0,96;1,33]	1,09 [0,88;1,35]	0,87 [0,56;1,30]		
LAM	O/E [d]	1 232 / 1 251,8	131 / 112,8	57 / 52,7	26 / 26,7	12 / 15,6	7 / 5,4	0,43	1,11 [0,75 ; 1,65]
	SIR [IC95%] [e]	0,98 [0,93;1,04]	1,16 [0,97;1,38]	1,08 [0,82;1,40]	0,97 [0,64;1,42]	0,77 [0,40;1,35]	1,29 [0,52;2,67]		

[a] e-t : écart-type de la moyenne
[b] moyenne pondérée par le nombre de PA de la période 1990-2009 des estimations faites à la mairie des communes de la classe d'exposition
[c] N : nombre de communes
[d] O : cas observés ; E : nombre de cas attendus de leucémie
[e] SIR : Standardized Incidence Ratio = O/E et intervalle de confiance à 95% estimé avec l'approximation de Byar
[f] p_h : p-value du test du KHI-2 d'hétérogénéité des SIR
[g] β : Coefficient dans la régression de Poisson du logarithme du SIR des leucémies sur la moyenne pondérée du radon de la classe considérée comme une variable quantitative ;exp(β) : facteur multiplicatif d'augmentation du SIR quand l'exposition au radon augmente de 100 Bq/m3

Tableau 16. Association entre l'exposition au radon estimée par cokrigeage et l'incidence des LAM par classe d'âge chez les enfants de 0 à 4 ans (N=709), de 5 à 9 ans (N=332) et de 10 à 14 ans (N=424) (France, 36 326 communes; 1990-2009)

		Classes d'exposition en Bq/m^3							
		<50	[50;100[	[100;150[	[150;200[	[200;250[	≥250	p_h [f]	exp(β) [g] IC95%
	moyenne ± e-t [a]	40,6 ± 1,2	68,5 ± 1,2	120,8 ± 1,1	169,2 ± 1,1	222,4 ± 1,1	340,3 ± 1,3		
	moyenne pondérée [b]	34,9	67,5	120,4	168,1	221,2	332,8		
	N [c]	6 266	19 441	6 740	1 999	784	1 096		
0-4 ans	O/E [d]	290 / 305,7	296 / 297,9	82 / 74,3	22 / 17,7	16 / 7,1	3 / 6,4	0,01	1,14 [0,97 ; 1,33]
	SIR [IC95%] [e]	0,95 [0,84;1,06]	0,99 [0,88;1,11]	1,1[0,88;1,37]	1,25 [0,78;1,89]	2,26[1,29;3,68]	0,47[0,09;1,37]		
5-9 ans	O/E [d]	137 / 137,1	142 / 143,2	32 / 36,1	13 / 8,8	2 / 3,5	6 / 3,2	0,36	1,11 [0,88 ; 1,39]
	SIR [IC95%] [e]	1,00 [0,84;1,18]	0,99 [0,84;1,17]	0,89 [0,61;1,25]	1,48 [0,79;2,53]	0,57 [0,06;2,05]	1,85 [0,68;4,03]		
10-14 ans	O/E [d]	197 / 172,0	168 / 184,5	47 / 47,1	4 / 11,5	6 / 4,6	2 / 4,2	0,04	0,76 [0,59 ; 0,98]
	SIR [IC95%] [e]	1,15 [0,99;1,32]	0,91 [0,78;1,06]	1,00 [0,73;1,33]	0,35 [0,09;0,89]	1,29 [0,47;2,81]	0,47 [0,05;1,71]		

[a] e-t : écart-type de la moyenne
[b] moyenne pondérée par le nombre de PA de la période 1990-2009 des estimations faites par cokrigeage à la mairie des communes de la classe d'exposition
[c] N : nombre de communes
[d] O : cas observés ; E : nombre de cas attendus de leucémie
[e] SIR : Standardized Incidence Ratio = O/E et intervalle de confiance à 95% estimé avec l'approximation de Byar
[f] p_h : p-value du test du KHI-2 d'hétérogénéité des SIR
[g] β : Coefficient dans la régression de Poisson du logarithme du SIR des leucémies sur la moyenne pondérée du radon de la classe considérée comme une variable quantitative ;exp(β) : facteur multiplicatif d'augmentation du SIR quand l'exposition au radon augmente de 100 Bq/m^3

Les résultats de l'étude cas-témoins GEOCAP sur l'association entre exposition au radon estimée par cokrigeage chez les 2 753 cas de LA et 30 000 témoins ne montrent pas non plus d'association de l'exposition au radon au diagnostic avec les leucémies de l'enfant (LA, LAL, LAM) ni globalement (Tableau 17), ni par classe d'âge.

Tableau 17. Association entre l'exposition au radon au diagnostic estimée par cokrigeage au domicile et l'incidence de toutes les leucémies aiguës (N=2 763), des LAL (N=2 283) et des LAM (N=418) chez les enfants de 0 à 14 ans dans l'étude cas-témoins GEOCAP (2002-2007)

		Classes d'exposition en Bq/m^3					
		<50	[50;100[	[100;150[	[150;200[	[200;250[	≥250
Témoins (N=30 000)		12 444	12 961	3 263	780	272	280
LA	O	1 120	1 228	298	70	30	17
	OR [IC95%] [a]	1,00	1,06 [0,98;1,16]	1,01 [0,89;1,16]	1,00 [0,78;1,29]	1,18 [0,80;1,74]	0,66 [0,40;1,09]
LAL	O	906	1 041	241	57	24	14
	OR [IC95%] [a]	1,00	1,11 [1,01;1,22]	1,01 [0,87;1,17]	1,00 [0,76;1,33]	1,16 [0,75;1,77]	0,67 [0,39;1,16]
LAM	O	193	157	49	12	5	2
	OR [IC95%] [a]	1,00	0,79 [0,64;0,98]	0,98 [0,72;1,35]	1,01 [0,56;1,81]	1,21 [0,49;2,98]	0,46 [0,11;1,87]

[a] Odds ratios (OR) et intervalles de confiance à 95% estimés par régression logistique ajustée sur l'âge en fin d'année (classes d'âge de 1 an)

Le Tableau 18 montre les SIR obtenus pour toutes les leucémies aiguës, les LAL et les LAM tous âges en fonction des classes d'exposition cumulée au radon estimée par cokrigeage. Aucune association n'est trouvée lorsque l'exposition est découpée en classes de 0,5 kBq/m^3 - année. Comme pour les analyses réalisées par classe d'âge d'exposition, les SIR sont tous proches de l'unité et les résultats des régressions de Poisson linéaires sur l'exposition cumulée sont non significatifs. Les résultats de l'étude cas-témoins vont dans le même sens.

Tableau 18. Association entre l'exposition cumulée au radon[a] et l'incidence de toutes les leucémies aiguës (N=9 056), des LAL (N=7 434) et des LAM (N=1 465) chez les enfants de 0 à 14 ans (France, 36 326 communes; 1990-2009)

		Classes d'exposition en kBq/m^3 – année							
		<0,5	[0,5;1[	[1;1,5[	[1,5;2[	[2;3[	≥3	p_h [g]	exp(β) [h] IC95%
	moyenne ± e-t [b]	0,25 ± 0,14	0,71 ± 0,14	1,21 ± 0,14	1,71 ± 0,14	2,37 ± 0,28	4,18 ± 1,18		
	moyenne pondérée [c]	0,24	0,69	1,20	1,71	2,36	4,05		
	N [d]	257 860	177 290	62 762	24 747	14 546	7 685		
LA	O/E [e]	6 536 / 6 549,9	1 857 / 1 811,4	423 / 445,8	140 / 147,4	75 / 73,5	25 / 28,1	0,69	0,99 [0,94;1,04]
	SIR [IC95%] [f]	1,00 [0,97;1,02]	1,03 [0,98;1,07]	0,95 [0,86;1,04]	0,95 [0,80;1,12]	1,02 [0,80;1,28]	0,89 [0,57;1,31]		
LAL	O/E [e]	5 392 / 5 431,4	1 521 / 1 461,7	335 / 349,0	102 / 113,8	66 / 56,7	18 / 21,5	0,25	1,00 [0,95;1,06]
	SIR [IC95%] [f]	0,99 [0,97;1,02]	1,04 [0,99;1,09]	0,96 [0,86;1,07]	0,90 [0,73;1,09]	1,16 [0,90;1,48]	0,84 [0,50;1,32]		
LAM	O/E [e]	1 035 / 1 010,0	300 / 315,7	81 / 87,6	34 / 30,4	8 / 15,3	7 / 6,0	0,31	0,93 [0,83;1,05]
	SIR [IC95%] [f]	1,02[0,96;1,09]	0,95[0,85;1,06]	0,92[0,73;1,15]	1,12[0,77;1,56]	0,52[0,23;1,03]	1,16[0,47;2,40]		

[a] Exposition cumulée au radon estimée à partir de l'estimation par cokrigeage de l'exposition au radon à la mairie de la commune de diagnostic
[b] e-t : écart-type de la moyenne /
[c] Moyenne pondérée par le nombre de PA de la période 1990-2009 des estimations faites à la mairie des communes de la classe d'exposition
[d] N : nombre de communes * âge
[e] O : cas observés ; E : nombre de cas attendus de leucémie
[f] SIR : Standardized Incidence Ratio = O/E et intervalle de confiance à 95% estimé avec l'approximation de Byar
[g] p_h : p-value du test du KHI-2 d'hétérogénéité des SIR
[h] β : Coefficient dans la régression de Poisson du logarithme du SIR des leucémies sur la moyenne pondérée du radon de la classe considérée comme une variable quantitative ; exp(β) : facteur multiplicatif d'augmentation du SIR quand l'exposition au radon augmente de 1 kBq/m3

B. RTC et leucémies de l'enfant

Le Tableau 19 donne les SIR obtenus dans l'étude d'incidence 1990-2009 par *zone d'emploi* (ZE) pour toutes les leucémies aiguës, les LAL et les LAM tous âges en fonction des classes d'exposition aux RTC estimée dans la *zone d'emploi* du diagnostic, avec un découpage en 6 classes de 20 nSv/h. Aucune association n'est observée, les SIR sont tous proches de l'unité. Les analyses par classe d'âge vont dans le même sens d'une absence d'association.

Après ajustement sur l'indicateur de défaveur, les résultats restent inchangés. Il en est de même, lorsque les analyses ont été restreintes aux 75% des ZE pour lesquelles l'estimation de l'exposition était la plus précise : c'est-à-dire en excluant les 63 *zones d'emploi* dont le coefficient de variation se trouvait dans le quartile supérieur de la distribution.

Les résultats de l'étude cas-témoins GEOCAP sur l'association entre exposition aux RTC chez les 2 512 cas de LA et 27 263 témoins habitant dans une *zone d'emploi* où l'estimation a pu être calculée ne montre aucun résultat significatif. Les résultats étaient également inchangés lorsqu'étaient exclus les cas et témoins résidant à moins de 50 m d'une ligne à très haute tension (225 kV ou 400 kV) ou ceux habitant à moins de 600 m tout comme pour l'exclusion des cas et témoins résidant dans un périmètre de 5 km ou 20 km autour d'une centrale nucléaire.

Les SIR obtenus pour toutes les leucémies aiguës, les LAL et les LAM tous âges en fonction des classes d'exposition cumulée ne montrent aucune association lorsque l'exposition est découpée en classes de 2 mSv. Comme pour les analyses réalisées par classe d'âge, les SIR sont tous proches de l'unité et les résultats de la régression de Poisson linéaires sur l'exposition cumulée n'indiquent pas de tendance.

Tableau 19. Association entre l'exposition aux RTC et l'incidence de toutes les leucémies aiguës N=8 219 (E=8 184), des LAL N=6 740 (E=6 718) et des LAM N=1 331 (E=1 324) chez les enfants de 0 à 14 ans (France, 249 zones d'emploi; 1990-2009)

		Classes d'exposition en nSv/h								
		<60	[60;70[	[70;80[	[80;90[	[90;100[	[100;110[	≥110	p_h [f]	exp(β) [g] IC95%
	moyenne ± e-t [a]	53,6 ± 1,1	65,2 ± 1,0	74,8 ± 1,0	84,4 ± 1,0	94,4 ± 1,0	104,5 ± 1,0	119,7 ± 1,1		
	moyenne pondérée [b]	53,7	65,7	74,7	84,2	94,4	105,4	119,1		
	N [c]	36	55	55	41	32	17	13		
LA	O/E [d]	1 275 / 1 299	2 630 / 2 687,8	1 751 / 1 704,7	966 / 940,8	968 / 923,4	414 / 403,8	215 / 224,9		
	SIR	0,98	0,98	1,03	1,03	1,05	1,03	0,96	0,37	1,01
	[IC95%] [e]	[0,93;1,04]	[0,94;1,02]	[0,98;1,08]	[0,96;1,09]	[0,98;1,12]	[0,93;1,13]	[0,83;1,09]		[1,00;1,02]
LAL	O/E [d]	1 035 / 1 066,8	2 136 / 2 205,1	1 470 / 1 399,8	786 / 772,5	800 / 758,1	337 / 331,4	176 / 184,6		
	SIR	0,97	0,97	1,05	1,02	1,06	1,02	0,95	0,14	1,01
	[IC95%] [e]	[0,91;1,03]	[0,93;1,01]	[1,00;1,11]	[0,95;1,09]	[0,98;1,13]	[0,91;1,13]	[0,82;1,10]		[1,00;1,03]
LAM	O/E [d]	214 / 209,8	443 / 436,0	259 / 275,3	162 / 152,0	146 / 149,4	70 / 65,4	37 / 36,4		
	SIR	1,02	1,02	0,94	1,07	0,98	1,07	1,02	0,90	1,00
	[IC95%] [e]	[0,89;1,17]	[0,92;1,12]	[0,83;1,06]	[0,91;1,24]	[0,83;1,15]	[0,83;1,35]	[0,71;1,40]		[0,97;1,04]

[a] e-t : écart-type de la moyenne
[b] moyenne pondérée par le nombre de PA de la période 1990-2009 des estimations des *zones d'emploi* de la classe d'exposition
[c] N : nombre de *zones d'emploi*
[d] O : cas observés ; E : nombre de cas attendus de leucémie
[e] SIR : Standardized Incidence Ratio = O/E et intervalle de confiance à 95% estimé avec l'approximation de Byar
[f] p_h : p-value du test du *KHI-2* d'hétérogénéité des SIR
[g] β : Coefficient dans la régression de Poisson du logarithme du SIR des leucémies sur la moyenne pondérée des RTC de la classe considérée comme une variable quantitative ; exp(β) : facteur multiplicatif d'augmentation du SIR quand l'exposition augmente de 10 nSv/h

Tableau 20. Association entre la dose à la moelle et l'incidence de toutes les leucémies aiguës N=8 219 (E=8 184), des LAL N=6 740 (E=6 718) et des LAM N=1 331 (E=1 324) chez les enfants de 0 à 14 ans (France, 249 zones d'emploi; 1990-2009)

		Classes d'exposition en mSv						
		<5	[5;10[	[10;15[	[15;20[	≥ 20	p_h [f]	exp(β) [g] IC95%
	moyenne ± e-t [a]	2,49 ± 1,43	7,48 ± 1,44	12,14 ± 1,4	16,74 ± 1,31	23,02 ± 2,56		
	moyenne pondérée [b]	2,51	7,48	11,92	16,63	22,6		
	N [c]	1 294	1 278	862	231	70		
LA	O/E [d]	4 505 / 4 466,8	2 381 / 2 384,8	1 074 / 1 055,6	219 / 231,7	40 / 45,7	0,73	0,98 [0,93 ; 1,04]
	SIR [IC95%] [e]	1,01 [0,98;1,04]	1,00 [0,96;1,04]	1,02 [0,96;1,08]	0,95 [0,82;1,08]	0,88 [0,63;1,19]		
LAL	O/E [d]	3 731 / 3 707,2	1 978 / 1 997,2	836 / 806,9	166 / 173,3	29 / 33,7	0,67	0,99 [0,93 ; 1,05]
	SIR [IC95%] [e]	1,01 [0,97;1,04]	0,99 [0,95;1,04]	1,04 [0,97;1,11]	0,96 [0,82;1,12]	0,86 [0,58;1,23]		
LAM	O/E [d]	693 / 687,7	365 / 347,2	212 / 225,5	50 / 53,1	11 / 10,8	0,74	0,96 [0,95 ; 1,09]
	SIR [IC95%] [e]	1,01 [0,93;1,09]	1,05 [0,95;1,16]	0,94 [0,82;1,08]	0,94 [0,70;1,24]	1,02 [0,51;1,83]		

[a] e-t : écart-type de la moyenne
[b] moyenne pondérée par le nombre de PA de la période 1990-2009 des estimations des *zones d'emploi* de la classe d'exposition
[c] N : nombre de *zones d'emploi*
[d] O : cas observés ; E : nombre de cas attendus de leucémie
[e] SIR : Standardized Incidence Ratio = O/E et intervalle de confiance à 95% estimé avec l'approximation de Byar
[f] p_h : p-value du test du *KHI-2* d'hétérogénéité des SIR
[g] β : Coefficient dans la régression de Poisson du logarithme du SIR des leucémies sur la moyenne pondérée de la classe considérée comme une variable quantitative ; exp(β) : facteur multiplicatif d'augmentation du SIR quand l'exposition augmente de 10 mSv

C. Radon, RTC et leucémies de l'enfant

Le Tableau 20 montre les SIR obtenus dans l'étude d'incidence 1990-2009 par ZE pour toutes les leucémies aiguës, les LAL et les LAM tous âges en fonction des classes d'exposition cumulée de dose à la moelle au niveau de la *zone d'emploi*, avec un découpage en 5 classes de 5 mSv. Aucune association n'est observée, les SIR sont tous proches de l'unité. Les analyses par classe d'âge vont dans le même sens d'une absence d'association. Après ajustement sur l'indicateur de défaveur, les résultats restaient inchangés comme après avoir pris en compte la dose reçue *in utero.*

D. Puissance de l'étude

Les résultats de l'évaluation de la puissance de nos études à l'échelle des communes indiquent une puissance de 79% sous les hypothèses alternatives de risque de modèle de référence ERR (UNSCEAR 2006).

IV. Discussion

Nous avons étudié l'association entre l'exposition aux radiations d'origine naturelle et l'incidence des leucémies de l'enfant par une étude d'incidence à l'échelle du territoire français pour la période 1990-2009 au cours de laquelle 9 056 cas de LA ont été enregistrés. L'exposition au radon a été estimée au niveau des communes. Celle des rayonnements gamma d'origine tellurique et cosmique au niveau de la *zone d'emploi*. Ceux-ci n'ont pas systématiquement été considérés dans les études semblables (Laurier et al., 2001). Une étude cas-témoins fondée sur des enregistrements et portant sur la période 2002-2007 a également été analysée. Aucune association n'a été observée que ce soit pour l'exposition au radon, pour l'exposition aux rayonnements gamma ou pour l'estimation de la dose à la moelle.

Parmi les études écologiques ayant étudié les mêmes expositions, onze des douze études ont montré des associations positives entre le niveau moyen de radon et l'incidence ou la mortalité de leucémie infantile; huit d'entre elles ont montré une association

significative. Les résultats concernant l'exposition aux radiations gamma donnaient des résultats plutôt mitigés. Notons que la moitié de ces études ont été faites au Royaume-Uni et que les autres ne sont pas toutes indépendantes les unes des autres parce que, comme en France (Evrard et al., 2005 ; Evrard et al., 2006a), les périodes et les populations se chevauchent. En France, l'étude écologique de l'équipe sur la période 1990-2001 a montré une association modérée mais significative entre l'exposition au radon domestique et l'incidence des leucémies aiguës myéloblastiques chez les enfants : l'incidence était de 19% supérieure dans les départements où l'exposition au radon était la plus forte (en moyenne, il y avait une différence de 100 Bq/m^3 entre le quintile d'exposition supérieur et le quintile inférieur pris comme référence) (Evrard et al., 2005 ; Evrard et al., 2006a). Aucune association n'a été trouvée pour les rayonnements gamma. Nous n'avons pas retrouvé d'association significative pour les LAM dans la présente étude pour la période 1990-2009 bien que l'association allait dans un sens positif et ce de façon plus nette pour la période 1990-1999.

Neuf études cas-témoins sur les leucémies de l'enfant et le radon ont été réalisées et présentent des résultats mitigés. Ces études diffèrent par la méthode d'estimation de l'exposition. Dans six d'entre elles, l'exposition a été évaluée à partir de mesures réalisées dans les maisons des enfants, alors qu'un modèle de prédiction était utilisé dans les trois autres études : au Danemark, au Royaume-Uni et en Suisse (Hauri et al., 2013a; Kendall et al., 2012; Raaschou-Nielsen et al., 2008). Ces méthodes ont chacune leurs avantages et leurs inconvénients. Les mesures permettent de tenir compte de la concentration dans la maison au cours de la période de mesure et également, contrairement à un modèle de prédiction, de refléter les habitudes d'aération réelles des habitants. Cependant cette méthode d'évaluation de l'exposition nécessite la coopération des familles et l'accès aux résidences (et peut-être également des anciennes maisons) des cas et des témoins, ce qui limite le taux de participation à ces études. Ces faibles taux de participation impliquent un risque de biais de sélection, en particulier si les taux de participation des cas et témoins diffèrent, ce qui peut affecter les estimations de risques dans une étude cas-témoins. Le

faible taux de participation qui diffère pour les cas et les témoins a été suggéré comme une explication de l'effet apparemment protecteur du radon sur le risque de leucémie dans l'étude anglaise UKCCS (UKCCS, 2002a; UKCCS, 2002b).

Une façon de garantir des taux de participation élevés est *d'estimer l'exposition dans les maisons par un modèle de prédiction et d'utiliser des bases de données d'enregistrement systématique de cas et de témoins* qu'on peut considérer comme exhaustives. Trois études nationales ont procédé de cette façon et rapportent des résultats différents. L'étude danoise (Raaschou-Nielsen et al., 2008) a rapporté une association significative entre l'exposition (cumulée) au radon et les LAL ce qui n'est pas le cas des études anglaise (Kendall et al., 2012) et suisse (Hauri et al., 2013a) (exposition périnatale au radon). L'étude anglaise a rapporté une association positive et significative des LA avec l'exposition (périnatale) aux rayonnements gamma.

Taille des études

Du point de vue de la taille, notre étude et l'étude anglaise (Kendall et al., 2012) sont les plus importantes avec plus de 9 000 cas enregistrés sur une période de respectivement 20 et 27 ans. L'étude danoise (Raaschou-Nielsen et al., 2008) bien que réalisée sur une plus longue période, enregistre un nombre de cas presque huit fois plus petit. L'étude suisse (Hauri et al., 2013a), est quant à elle, basée sur seulement 283 cas de LA.

Méthodes d'estimation des expositions.

Dans l'étude suisse (Hauri et al., 2013a), l'estimation au radon est prédite par un modèle incluant des données géologiques et des caractéristiques du logement qui expliquent 20% de la variance de la concentration en radon à l'intérieur des habitations. L'information sur les caractéristiques du logement des sujets de l'étude est connue grâce aux données de la *Swiss national cohort*. Dans l'étude anglaise (Kendall et al., 2012), deux sources d'estimation de concentration du radon pour les maisons des sujets de l'étude ont été utilisées: l'exposition moyenne du *County District* calculée à partir de 2 283 mesures issues d'une campagne nationale (Wrixon and Board, 1988) et une carte prédictive du radon basée sur les résultats d'environ 400 000 mesures groupées par carrés et par unités géologiques

(Appleton and Miles, 2010; Miles and Appleton, 2005). Entre 34% et 40% de la variance de la concentration en radon est expliquée par cette carte selon le type de géologie (Appleton and Miles, 2010). Dans l'étude conduite au Danemark, 40% de la variabilité de près de 3 000 mesures intérieures de radon a été expliquée par le modèle (Andersen et al., 2007). Dans ce modèle, les variables explicatives les plus importantes étaient le type de maison et la géologie. La géologie expliquait une partie de la variabilité similaire à celle expliquée par la géologie dans notre modèle utilisé pour déterminer le score d'exposition alors que la variabilité expliquée par le logement était supérieure à celle expliquée dans la présente étude. L'information sur les caractéristiques du logement des sujets de l'étude est connue grâce aux données du *Building and Dwelling Register* danois. Les études suisse et danoise ont ainsi l'avantage de disposer des caractéristiques du logement où résident les sujets de l'étude. En France et au Royaume-Uni, ces caractéristiques expliquent une part plus faible de la variabilité du radon (cf Partie V. V) (Appleton and Miles, 2010). L'exposition moyenne la moins élevée est celle du Royaume-Uni avec une moyenne géométrique de 16,4 Bq/m^3 (± 2,0) et une moyenne arithmétique de 21,3 Bq/m^3 (± 22,6), malgré la présence de valeurs élevées : les valeurs vont de 1,2 à 692 Bq/m^3. Au Danemark, bien que plus élevée en moyenne (48 Bq/m^3), l'exposition des sujets est moins variable avec une exposition allant de 4 à 254 Bq/m^3. En Suisse la moyenne arithmétique de l'exposition de la population est de 86 Bq/m^3 avec une étendue de 0,7 à 490 Bq/m^3. En France, la moyenne arithmétique de l'exposition des témoins GEOCAP est 67,8 Bq/m^3 avec des valeurs allant de 12,8 à 801,3 Bq/m^3. Il ne semble donc pas que les différences d'exposition entre les pays expliquent les différences de résultats.

L'exposition aux rayonnements gamma a été prise en compte dans l'étude suisse mais uniquement comme variable d'ajustement. Elle a été prise en compte de façon plus précise dans l'étude de Kendall et al. au Royaume-Uni (Kendall et al., 2012). L'estimation basée sur un ensemble de 2 283 mesures a été estimée par *County District* ce qui représente en moyenne une population de 100 000 habitants (Kendall et al., 2006). C'est un peu moins que la population d'une *zone d'emploi* en France où l'estimation est basée sur plus de 14 000

mesures au total. Par ailleurs au Royaume-Uni, en raison de l'appariement géographique des cas et des témoins au niveau du registre de naissance, les expositions de seulement 52 % des cas et des témoins étaient différentes (95% pour l'analyse du radon). En France, l'estimation de l'exposition aux RTC a pu être estimée pour 249 des 304 *zones d'emploi*. Cela représente 90% du nombre de cas attendus en France entière. La variabilité de l'exposition des sujets est semblable dans les deux études : dans l'étude anglaise celle-ci varie de 38,1 nGy/h à 159,7 nGy/h avec une moyenne arithmétique de 94,7 nGy/h (± 15,6) tandis que l'exposition estimée pour les *zones d'emploi* en France donne une moyenne arithmétique brute de 78 nSv/h avec une étendue de 46,2 nSv/h à 140,1 nSv/h. Les résultats partiellement contradictoires de ces deux études pourraient, notamment, provenir du fait que l'estimation des expositions pourrait être améliorée en augmentant le nombre de mesures.

Expositions périnatale, cumulée, au diagnostic et histoires résidentielles

Un autre aspect méthodologique important concerne les résidences incluses dans l'évaluation de l'exposition. Comme les enfants peuvent avoir vécu dans plus d'un foyer jusqu'au moment du diagnostic, le choix idéal consiste à inclure toutes les adresses dans lesquelles les cas et les témoins ont vécu pendant l'enfance afin de calculer l'exposition cumulée de la naissance jusqu'au moment du diagnostic ou de l'inclusion dans l'étude pour les témoins. Cela a été fait au Danemark grâce aux données du *Danish Central Population Registry*. Les sujets de l'étude avaient résidé à 2 adresses différentes en moyenne. Dans l'étude anglaise, l'estimation de l'exposition cumulée a été faite à partir de l'exposition à la résidence de la mère à la naissance. 50% des cas de cette étude (l'information n'étant pas disponible pour les témoins) n'avaient jamais déménagé et 20% supplémentaires avaient déménagé à moins de 2 km. Dans notre étude, nous ne disposions pas de l'histoire résidentielle des sujets mais de la commune ou de l'adresse au diagnostic. L'exposition cumulée a été calculée à partir de l'âge et de l'exposition au diagnostic après avoir vérifié que la mobilité des sujets n'était pas trop grande et que les expositions aux communes de résidence à la naissance et au diagnostic étaient très corrélées dans les études cas-témoins Escale et Estelle.

Approche écologique, approche individuelle

L'étude cas-témoins individuelle GEOCAP nous ouvrait *a priori* la perspective de pouvoir mettre en relation au niveau individuel une information d'exposition liée à l'adresse du domicile et une information sur le statut cas ou témoin. Toutefois la nature des informations d'exposition dont nous disposions et la structure de leur variabilité spatiale nous ont conduits à considérer nos informations d'exposition au niveau d'unités géographiques et donc à adopter une approche écologique :

- la corrélation égale à 0,99 entre l'estimation cokrigée au radon à l'adresse du domicile et à celle de la mairie de la commune de résidence a bien montré, comme l'existence d'un *effet pépite* majeur (60% de la variabilité) que le cokrigeage ne pouvait pas donner une estimation individuelle de l'exposition à une certaine adresse, mais seulement une information d'exposition moyenne au niveau d'une commune,

- pour les expositions aux radiations gamma telluriques ou cosmiques, les données dont nous disposions n'autorisaient qu'une estimation moyenne par *zone d'emploi*.

D'individuelle *a priori*, notre échelle d'étude est donc devenue écologique *de facto*, nous autorisant d'ailleurs à bénéficier de 20 années d'observations (1990-2009) plutôt que des seules 6 années de l'étude GEOCAP (2002-2007).

Il est bien établi que les résultats des études de corrélation écologique fondées sur des indicateurs d'exposition et de santé agrégés au niveau de groupes ne sont pas nécessairement cohérents avec ceux des études épidémiologiques réalisées au niveau individuel, cas-témoins et cohortes, notamment. La notion de corrélation établie au niveau écologique, peut avoir un intérêt en soi quand les phénomènes que l'on étudie ne se définissent qu'au niveau de populations et pas au niveau individuel, comme par exemple lorsqu'on s'intéresse au lien entre un taux de couverture vaccinale et la probabilité de survenue d'épidémies. C'est l'extension au niveau individuel d'une association observée au niveau écologique qui pose question.

L'étude écologique que nous avons développée a des caractéristiques bien spécifiques qui en justifient l'intérêt :

• elle vise à éprouver une hypothèse légitimée par des connaissances préalables solides portant sur l'association entre des expositions bien spécifiques, les radiations d'origine naturelle et des pathologies bien spécifiques dont on connaît les facteurs de risques principaux bien établis et ceux qui font l'objet d'hypothèses étayées (ici les leucémies de l'enfant et leurs liens avec les radiations ionisantes) et ne portent donc pas sur des ensembles très vastes et divers d'expositions et de pathologies dont on décrit les multiples corrélations écologiques sans hypothèse spécifique structurante sous-jacente.

• elle a été réalisée à une échelle géographique suffisamment étendue et sur une période de temps suffisamment longue pour disposer d'indicateurs d'incidence fondés sur de nombreux cas, d'indicateurs d'exposition moyenne fiables, de contrastes d'exposition marqués garantissant ensemble une puissance statistique *a priori* élevée sous des hypothèses de force d'association écologique argumentée.

• elle s'inscrit dans un contexte épidémiologique, où :

- le seul facteur de risque majeur et connu des leucémies de l'enfant susceptible d'interférer avec l'étude du lien écologique entre radiations naturelles et leucémies de l'enfant est l'âge,

- les autres facteurs de risque individuels établis (gènes IKZF1 et ARID5B, trisomie 21, radiations ionisantes à fortes doses, certaines chimiothérapies anticancéreuses) ont une faible vraisemblance d'interférer avec les relations entre radiations naturelles et leucémies de l'enfant aussi bien au niveau individuel qu'au niveau écologique.

L'approche écologique que nous avons adoptée constitue donc une base solide pour tester l'hypothèse que l'exposition domestique aux radiations naturelles est associée à un accroissement du taux d'incidence des leucémies pendant l'enfance et pour en quantifier l'importance. L'étude cas-témoins GEOCAP permet, par ailleurs, de prendre en compte certains facteurs environnementaux de variabilité au niveau individuel si nécessaire.

Prise en compte des disparités spatiales socio-économiques.

Dans notre étude, nous avons considéré un indicateur socio économique basé sur des données acquises au niveau de la commune au moment du recensement de 1999 : le taux

de chômage et d'ouvriers dans la population active, le taux de bacheliers et le revenu fiscal médian. Chacune de ces variables représente une dimension fondamentale du niveau socio-économique et son sens est homogène sur l'ensemble du pays (Rey et al., 2009).

La prise en compte de cet indicateur de défaveur au niveau communal n'a pas modifié nos résultats ni qualitativement ni quantitativement.

Une autre approche aurait pu être celle de Lalloue et al. qui ont réalisé une analyse en composantes principales sur 48 variables et construit des indices au niveau de l'IRIS dans plusieurs agglomérations (Strasbourg, Lille, Lyon, Marseille) (Lalloue et al., 2013). Au Royaume-Uni l'indice de Carstairs tient compte du chômage, du fait de posséder une voiture ou non, du surpeuplement du logement et de la classe sociale élevée ou non. Il existe un autre indice, l'indice de Townsend dans lequel le fait d'être propriétaire de son logement remplace l'indicateur de classe sociale. Pornet et al. ont utilisé une enquête sur les revenus et conditions de vie pour construire un indicateur de déprivation sociale à visée européenne EDI (European Deprivation Index) (Pornet et al., 2012). Concernant le risque de leucémie, les résultats restent divergents. D'après une revue de 47 études (Poole et al., 2006): la moitié des études rapportait un risque non significatif pour les leucémies et dans l'autre moitié, autant d'études rapportaient un lien positif que négatif. D'après une revue plus récente de 7 études (Adam et al., 2008), 4 études rapportent une association positive, 2 une association négative et la dernière étude une association non significative. Ces associations étaient faibles quand elles existaient. Carozza et al., dans une étude cas-témoins sans interview, n'ont pas rapporté d'association avec le niveau d'éducation des parents (Carozza et al., 2010). Youlden et al., dans une étude écologique Australienne, n'a pas montré d'association significative avec un indice composite de niveau socio-économique (Youlden et al., 2012). Dans l'étude de Kroll (Kroll et al., 2011), une association positive a été mise en évidence avec le niveau socio-économique au Royaume-Uni.

Le manque de données épidémiologiques convaincantes pour mettre en évidence une augmentation de l'incidence des leucémies infantiles due aux radiations naturelles de fond peut être lié à un certain nombre de raisons. Le nombre de cas radio-induits par les

radiations de fond pourrait être trop petit pour être détecté - si, par exemple le transfert de risque est absolu ou si les modèles dose-réponse donnés par les modèles de risque standards surestiment le risque de leucémie à faibles doses. D'autre part, la puissance statistique de la plupart des études qui ont été réalisées peut être insuffisante pour identifier un risque dû au rayonnement de fond qui serait, de fait, différent de zéro. Un nombre trop petit de cas inclus dans une étude est une cause évidente de faible puissance statistique mais un aspect qui peut aussi être important est la variation géographique ou inter-individuelle de l'exposition. Les corrélations possibles entre l'exposition aux rayonnements et d'autres facteurs qui influent sur la leucémie, en particulier dans les unités pour lesquelles l'exposition a été moyennée, pourraient aussi dissimuler une association réelle. Aucune hypothèse fortement étayée ne permet à cette étape d'identifier un ou des facteurs de risque, qui auraient des variations spatiales liées à celles des radiations ionisantes d'origine naturelle et qui pourraient avoir un effet de confusion négative avec ces expositions. L'existence d'un tel facteur ne peut pas, pour autant, être exclue *a priori*. On peut penser, par exemple, à des facteurs de risque environnementaux suspectés tels que l'exposition aux UV ou au benzène dû au trafic routier. Notons aussi que les zones à fort potentiel émissif de gaz radon sont plutôt rurales et de faible densité démographique. La densité de radiologues y est sans doute plus faible, cette profession médicale ayant tendance à se concentrer dans les zones urbaines à revenu élevé. On peut donc faire l'hypothèse que la fréquence des actes de radiologie y est relativement faible, et ainsi donc que l'exposition *in utero* ou à des âges très jeunes y est également, en moyenne, plus faible que la moyenne nationale. Le rôle possible de ces facteurs est pour certains, benzène dû au trafic routier et UV, actuellement étudié dans l'équipe.

Si la variation de la dose à la moelle entre les zones ou les individus est faible, alors il sera difficile de discerner un excès de risque résultant des rayonnements de fond. En effet, la proportion de cas causés par le rayonnement de fond pourrait être importante et ne pas être décelée par une grande étude, car il existe une variation insuffisante du risque pour permettre sa manifestation. Little et al. (Little et al., 2010) ont fait remarquer que dans les

études épidémiologiques sur le cancer et les rayonnements ionisants, la puissance statistique pourrait être beaucoup plus influencée par des différences dans la dose moyenne que par le nombre de cas. Les grandes différences dans les niveaux d'exposition de notre population à l'étude comme les premiers résultats de notre étude de puissance vont dans le sens d'une puissance élevée. En termes de puissance statistique et concernant les erreurs de classification de l'exposition, notre étude peut essentiellement être sujette à une erreur du type Berkson car nous avons utilisé un modèle de prédiction (Raaschou-Nielsen, 2008; Steenland et al., 2000). Contrairement aux erreurs de mesures individuelles, ce type d'erreur ne peut pas biaiser les estimations des associations (Steenland et al., 2000). L'échelle écologique réduit la variabilité totale des expositions et donc la puissance que pourraient atteindre ces études individuelles mais la validité de celles-ci est sérieusement fragilisée par des biais de participation importants.

De nombreuses recherches ont été effectuées sur la radioactivité naturelle et ses effets sur la santé avec pour objectif l'évaluation du risque attribuable à l'exposition. Plusieurs auteurs ont évalué la fraction de la leucémie infantile qui pourrait être attribuée à l'exposition au radon à partir des résultats observés dans les études épidémiologiques. A partir de l'étude écologique menée en France, Evrard et al estiment que 5,4% (IC95% [0,01% ; 11,3%]) des cas de LA de la petite enfance pourraient être attribués au radon (Evrard et al., 2005). De leur étude cas-témoins menée au Danemark (concentration moyenne annuelle de radon pondérée par la population 59 Bq/m^3), Raaschou -Nielsen et al. estiment qu'environ 9% des cas de LAL pourraient être attribués à l'exposition au radon.

Des évaluations classiques du risque attribuable ont également été réalisées, en utilisant les relations dose/risque de la littérature (généralement obtenus à partir des résultats du suivi des survivants d'Hiroshima et de Nagasaki) et les estimations de la dose équivalente à la moelle. Il a été suggéré, en 2004, que près de 6% des cas de leucémie infantile pouvaient être attribués au radon au Royaume-Uni (Kendall and Smith, 2005). Dans un travail plus récent, l'estimation de dose équivalente moyenne due aux sources naturelles d'exposition en Grande-Bretagne a été estimée à environ 1,3 mSv par an pour les enfants de

moins de 15 ans (Little et al., 2010). En utilisant ces estimations de dose, les auteurs ont estimé que l'exposition naturelle pouvait représenter 15 à 20% de tous les cas de leucémie infantile en Grande-Bretagne, bien que les incertitudes associées à cette estimation soient considérables (Little et al., 2009; Wakeford et al., 2009). Le radon contribuant à environ 10% de la dose à la moelle, la fraction attribuable à l'exposition au radon peut être estimée à environ 2% (0,5 à 3% selon le modèle). D'après ces calculs, les auteurs avaient conclu qu'il était peu probable que le radon puisse influencer le risque de leucémie infantile en Grande-Bretagne de façon détectable (à l'égard des autres sources de rayonnement de fond) (Wakeford et al., 2010). Néanmoins, la situation peut être différente dans d'autres pays avec des concentrations en radon plus élevées comme en France. Une évaluation de risque a récemment été conduite par l'IRSN en France (Laurent et al., 2013). Les doses cumulées à la moelle osseuse pendant l'enfance dues au radon, aux rayons gamma telluriques et cosmiques ont été estimées à 4,4, 7,5 et 4,3 mSv respectivement. Les excès de cas attribuables associés à ces sources de rayonnement naturel ont été estimés selon les modèles de l'UNSCEAR 2006 (UNSCEAR, 2006) et sont de 20% (intervalle de crédibilité 95%: 0-68%) et 4% (intervalle de crédibilité 95%: 0-11%) selon les modèles avec excès relatif et risque absolu, respectivement. Les grands intervalles de crédibilité, ainsi que les différentes estimations ponctuelles obtenues dans ces deux modèles, mettent en évidence les incertitudes dans les prévisions du risque de leucémie lié aux radiations, notamment liées au relativement faible nombre de cas de leucémies de l'enfant, dix cas en tout, observés dans la LSS. Au Royaume-Uni, l'association observée entre leucémie de l'enfant et dose à la moelle est à un niveau compatible avec les prédictions des modèles classiques (Kendall et al., 2012) suggérant que les estimations du risque de leucémie issues de données de la LSS sont largement applicables aux débits de dose très faibles reçus par les radiations d'origine naturelle.

L'épidémiologie fournit les principales bases de l'estimation des effets stochastiques des expositions aux radiations ionisantes pour l'homme. Les résultats des différentes études épidémiologiques sont régulièrement recensés par des comités scientifiques internationaux

et les modèles de risque réévalués. Les dernières publications ont été réalisées par l'UNSCEAR (2006) et le BEIR VII (2007). Comme souligné par l'UNSCEAR (2008) et d'autres auteurs (Little, 2001a; Little, 2001b; Little, 2008), en général, les risques relatifs par unité de dose chez les survivants des bombardements atomiques au Japon sont supérieurs à ceux des études des individus médicalement exposés à des doses élevées. Bien que d'autres facteurs puissent y contribuer, la question est souvent posée d'une relation exposition-risque différente selon les caractéristiques de la population. La même question se pose au sujet des conséquences des expositions subies durant l'enfance. Du fait des difficultés méthodologiques inhérentes à l'étude de l'effet d'une exposition dans l'enfance sur un risque influencé par de nombreuses autres expositions et qui se manifeste surtout plusieurs dizaines d'années après, il existe peu d'informations permettant de se positionner solidement.

Bien que les études sur les survivants japonais des bombardements de 1945 soient impressionnantes dans le détail des informations sur les risques radiologiques qu'elles fournissent, elles ne peuvent pas donner d'information directe sur tous les aspects des risques induits par les radiations. Ces survivants ont reçu des doses délivrées brièvement, principalement de rayonnement gamma et de façon externe, et quelques-unes des estimations de la dose restent incertaines. La population exposée souffrait de malnutrition à la fin d'une longue guerre, et pour entrer dans la LSS les survivants devaient avoir vécu jusqu'en Octobre 1950 dans des conditions qui étaient loin d'être idéales. Ceci soulève la possibilité que ceux qui entrent dans les études épidémiologiques étaient des individus plus forts, non représentatifs de la population générale en termes de risques de cancer radio-induit. Se pose aussi la question de savoir comment l'excès de risque relatif ou l'excès de risque absolu sont transférables entre les populations? La population japonaise exposée en 1945 avec notamment son propre risque de cancer de fond (par exemple, un risque relativement élevé de cancer de l'estomac, mais un risque relativement faible de cancer du sein chez la femme) a des caractéristiques différentes de la population actuelle de l'Europe de l'Ouest, avec un risque de cancer différent (par exemple un risque relativement faible de

cancer de l'estomac, mais un risque relativement élevé de cancer du sein chez la femme). L'incidence de la leucémie infantile dans les années 1950 était sensiblement plus faible au Japon qu'en Grande-Bretagne ou aux Etats-Unis (ou dans le Japon contemporain), de sorte que de savoir comment l'excès de risque radio-induit de leucémie infantile chez les survivants des bombardements atomiques se rapporte à celui d'autres populations reste une question pertinente.

Si les résultats des analyses épidémiologiques plus récentes apportent un faisceau d'arguments en faveur de la cancérogénicité des expositions aux radiations à des niveaux et des débits de dose élevés et brefs, ils ne peuvent être directement informatifs pour les expositions les plus courantes qui sont à faible niveau et débit de dose faible et continu. Pour quantifier le risque en population générale, il est également nécessaire de se poser la question de l'existence d'une relation sans seuil et de la linéarité. La forme de la relation semble plutôt linéaire quadratique (UNSCEAR) mais en l'absence de preuve expérimentale, il existe toujours une discussion d'experts sur l'existence ou l'absence d'effet seuil. A de très faibles niveaux de doses la quantification de la relation dose-réponse est encore incertaine car fondée sur des observations directes encore insuffisamment informatives.

Conclusion

Les observations danoises (Raaschou-Nielsen et al., 2008) mettent en évidence une association entre exposition au radon et leucémies de l'enfant. L'étude britannique (Kendall et al., 2012) ne montre pas de telle association avec le radon mais une association avec l'exposition aux rayonnements gamma est observée. Nos observations ne mettent pas en évidence de telles associations. Pourtant, du fait de l'importance du nombre de cas observés, de l'exhaustivité de leur enregistrement, de l'importance des contrastes d'expositions aux radiations ionisantes d'origine naturelle existantes entre communes ou zones d'emploi, nous avons estimé que nos études auraient une très bonne puissance pour mettre en évidence de telles associations, à des niveaux quantitatifs qui seraient compatibles avec les modèles de risque de référence de l'UNSCEAR (2006). Ces estimations de puissance seraient valides sauf si :

- les expositions moyennes des unités géographiques utilisées (communes et zones d'emploi) étaient grossièrement mésestimées. Ceci nous semble peu plausible même si nos estimations indirectes des expositions périnatales pourraient être améliorées.

- un éventuel effet des variations géographiques des expositions aux radiations ionisantes d'origine naturelle en France était négativement confondu avec celles d'un ou plusieurs facteurs de risque de ces leucémies de l'enfant, hypothèse que nous ne devons pas exclure mais qui ne peut être étayée qu'en identifiant et prenant en compte des facteurs susceptibles d'avoir un tel effet de confusion.

La question posée nous semble suffisamment importante du point de vue de la radioprotection pour que soient réalisées des études analogues dans d'autres pays et des études complémentaires au Danemark, au Royaume-Uni, en Suisse et en France afin d'apporter de nouveaux éléments d'appréciation épidémiologique directe de la validité des modèles de risque de référence en exposition chronique, à faible dose dans diverses populations.

Partie VII Conclusion générale et perspectives

I. Conclusion générale

L'objectif de cette thèse était d'étudier les variations géographiques de l'incidence des leucémies de l'enfant et l'association avec l'exposition aux radiations ionisantes d'origine naturelle.

Un premier travail a consisté à étudier la répartition spatiale de l'incidence des leucémies de l'enfant au niveau des 1 916 *bassins de vie* définis par l'INSEE. Des méthodes de détection de cluster ont été appliquées sur les 7 675 cas de leucémies de l'enfant diagnostiqués au cours de la période 1990-2006 afin d'identifier les zones potentiellement associées à un plus fort risque de leucémies aiguës de l'enfant. Cette étude, qui portait sur l'ensemble de la France métropolitaine n'a pas mis en évidence d'hétérogénéité spatiale des taux d'incidence des LA de l'enfant au cours de la période 1990-2006 au niveau des BV. Cependant, quelques clusters spatiaux ont été identifiés dans des lieux et périodes spécifiques. Bien que les niveaux de significativité des clusters ne soutiennent pas fortement l'existence de facteurs de risque localisés, les clusters peuvent montrer un léger impact de facteurs de risque partagés à l'échelle des BV, y compris dans le contexte d'expositions environnementales, qui nécessitent une enquête plus approfondie.

Pour tester l'hypothèse de l'existence d'une association entre l'exposition environnementale aux radiations ionisantes d'origine naturelle et l'incidence des leucémies de l'enfant, une étude d'incidence basée sur les 9 056 cas de LA de la période 1990-2009 a été réalisée. Cette étude a été complétée par une étude cas-témoins en population fondée sur les 2 763 cas de LA enregistrés sur la période 2002-2007 et un ensemble témoin de 30 000 sujets, construit en lien avec l'INSEE, constituant un échantillon contemporain représentatif de la population pédiatrique française. Dans cette approche, la géolocalisation des adresses des cas et des témoins ainsi que la géolocalisation des sources d'exposition et leur caractérisation permet de définir les critères de l'intensité d'exposition aux facteurs d'intérêt et de les mettre en relation avec le statut cas *vs* témoins des sujets.

Notre étude n'a pas montré d'association entre les leucémies de l'enfant et les expositions d'origine naturelle au radon et aux rayonnements gamma d'origine tellurique et cosmique estimés au diagnostic et de façon cumulée pendant l'enfance. Elle avait cependant une bonne puissance pour mettre en évidence les risques attendus d'après les modèles de risque actuels (BEIR, UNSCEAR) issus des études sur les risques observés à forte dose. Si ce travail n'étaye pas l'hypothèse qu'il existerait une association de ce niveau d'intensité entre les expositions aux rayonnements ionisants d'origine naturelle observés en France et l'incidence des leucémies de l'enfant, il apporte des informations suffisamment solides pour considérer que la question initiale reste posée et mérite que soient développées des études complémentaires plus approfondies aussi bien en France que dans d'autres pays.

II. Perspectives

Concernant l'estimation de l'exposition au radon, la réalisation de mesures complémentaires est en cours de discussion entre l'IRSN et l'Institut de veille sanitaire (InVS), en particulier pour améliorer la connaissance de l'exposition dans les zones les moins bien couvertes et présentant une variabilité locale importante, et pour évaluer la stabilité de l'exposition en fonction de l'évolution de l'habitat. Des données supplémentaires sur les rayonnements gamma pourraient contribuer à l'amélioration de la précision de l'estimation de l'exposition. Des mesures réalisées avec des dosimètres radio photo luminescents (RPL) remplaçant les TLD, plus homogènes entre eux et avec possibilité de relecture, des données de mesures réalisées par des GPS véhiculés pourraient y contribuer tout comme les informations d'exposition médicale recueillies par l'IRSN (IRSN/BUDI) auprès d'environ 25 000 travailleurs avec dosimètre dans des cabinets médicaux et de dentistes.

De manière plus générale, dans le cadre du premier Plan National Santé Environnement, un premier plan national d'actions 2005-2008 relatif à la gestion du risque lié au radon avait été élaboré par l'ASN, en 2005, en collaboration avec la Direction de l'habitat, de l'urbanisme et des paysages (DHUP), l'IRSN, l'InVS et le Centre scientifique et technique

du bâtiment (CSTB). Cette dynamique s'est poursuivie et amplifiée dans le second plan national d'actions 2011-2015 pour la gestion du risque lié au radon. Elaboré en collaboration entre l'ASN, les ministères chargés de la santé, de la construction et du travail et les organismes partenaires (IRSN, InVS, CSTB et le conseil régional du Limousin), ce nouveau plan inscrit en première priorité la réduction des expositions dans l'habitat existant et l'application de nouvelles règles de construction dans les bâtiments neufs afin de garantir un niveau faible d'exposition des personnes. De façon encore plus générale : le radon, un gaz naturel inerte et radioactif, mais dépourvu d'odeur, de couleur ou de goût est reconnu cancérigène pour le poumon (IARC, 1988). Pour autant, selon le Baromètre cancer 2010 (Beck et al., 2013), le radon reste encore largement méconnu du grand public, y compris dans les départements identifiés comme concernés par ce risque. Parmi les personnes âgées de 15 à 75 ans interrogées dans la population française, seule une sur cinq savait que le radon était un gaz naturel provenant du sol.

Concernant les données de leucémies, un nouveau programme, HOPE-EPI, est développé par Jacqueline Clavel. L'objectif est de proposer une infrastructure de recherche nationale mutualisant des efforts de collecte de données et de validation d'informations et d'évènements de santé au service d'un programme de recherche épidémiologique sur les facteurs de risque environnementaux et génétiques des cancers de l'enfant, les disparités de prise en charge et de leurs déterminants, les effets iatrogènes des traitements et plus généralement l'état de santé à court, moyen et long terme des personnes ayant eu un cancer dans l'enfance et les déterminants de leur qualité de vie.

Dans ce programme, il est prévu que les histoires résidentielles des cas soient enregistrées de manière systématique. Ainsi, l'exposition à un moment précis (par exemple au moment de la conception) pourra être étudiée, tout comme l'exposition cumulée entre la conception et le diagnostic. Une première étape, avant le recueil de l'histoire résidentielle des cas du RNHE, pourra être l'étude de l'exposition à la commune de naissance. Compte tenu de l'importance généralement accordée dans la littérature au rôle des expositions prénatales aux radiations ionisantes dans la survenue des leucémies de l'enfant, il serait

intéressant de vérifier si, comme nous le pensons, les expositions aux radiations à l'habitation à la naissance et à la commune de naissance (dont nous disposons pour tous les cas) sont plus liées entre elles que les expositions des habitations à la naissance et au diagnostic.

Enfin, le recueil des observations de la nouvelle enquête cas-témoins par questionnaire (Estelle) qui vient de s'achever comporte le relevé de l'histoire résidentielle des sujets, depuis leur naissance, qui permettra une estimation de l'histoire de l'exposition d'intérêt, comme le radon, de la naissance au diagnostic et aussi de mettre en corrélation les différents types de trajectoires résidentielles avec d'éventuelles particularités des sujets et/ou de leur parents.

Concernant l'étude GEOCAP, une étude de faisabilité du recueil exhaustif de l'adresse à la naissance est en cours : « GEOCAP-Past » sera basée sur les résidences à la naissance pour cibler les expositions environnementales de la période prénatale. Par ailleurs, le RNHE étant exhaustif pour les années 2008 à 2010 depuis la fin d'année 2012, le tirage d'un nouvel échantillon témoin pour ces mêmes années a été mis en place avec l'INSEE et depuis géolocalisé. D'autres volets de l'exposition environnementale en lien avec l'incidence des cancers de l'enfant sont et vont être étudiés grâce au programme GEOCAP et notamment les expositions :

- aux radiations UV et aux pesticides liés aux activités agricoles,
- au benzène dû au trafic routier et également à la proximité aux stations-services.
- aux champs électriques et magnétiques à extrêmement basse fréquence (CEM-EBF) liés à la proximité des lignes à haute tension (étape suivante de l'analyse réalisée par l'équipe sur la distance) qui porte sur l'exposition de chaque domicile aux CEM-EBF générés par les lignes à haute tension. Rte a développé une méthode d'estimation des expositions aux CEM-EBF à proximité des lignes aériennes à haute tension. Cette méthode prend principalement en compte la distance à la ligne et au pylône, la hauteur des pylônes, la nature des câbles, la géographie de la ligne et le courant moyen annuel dans cette ligne (Bessou et al., 2013).

Enfin, une mesure définie dans le Plan cancer 2009-2013 visait à « améliorer la prise en charge et à lancer un programme d'actions spécifiques vis-à-vis des adolescents atteints de cancer ». Cette mesure a été déclinée au niveau du programme partenarial Francim-Institut de veille sanitaire (InVS)-Institut national du cancer (INCa)-Hospices civils de Lyon par la surveillance des 15-24 ans par les registres généraux et spécialisés du réseau Francim, ce qui permet de fournir en routine des données d'épidémiologie descriptives d'incidence et de survie. L'extension de l'âge d'inclusion jusqu'à 17 ans révolus à partir du 1er janvier 2011 dans les registres nationaux pédiatriques : le Registre national des tumeurs solides de l'enfant (RNTSE) et le RNHE constitue la deuxième mesure. Le RNTSE enregistre les tumeurs de l'enfant depuis 2000 (950 cas par an dont 350 tumeurs du système nerveux central). Le RNHE et le RNTSE ont fonctionné de façon étroitement coordonnée depuis leur mise en place et partagent maintenant la même base de données et pourront donc être utilisés pour observer des périodes d'enregistrement plus longues, des étendues d'âge plus importantes et différents types de cancers de l'enfant et de l'adolescent.

Partie VIII Bibliographie

Adam M, Rebholz CE, Egger M, Zwahlen M, Kuehni CE. Childhood leukaemia and socioeconomic status: what is the evidence? Radiat Prot Dosimetry 2008; 132: 246-54.

Ahlbom A, Day N, Feychting M, Roman E, Skinner J, Dockerty J, et al. A pooled analysis of magnetic fields and childhood leukaemia. Br J Cancer 2000; 83: 692-8.

Alexander FE. Radon and Leukaemia. Lancet 1990; 335: 1336-1337.

Amigou A, Sermage-Faure C, Orsi L, Leverger G, Baruchel A, Bertrand Y, et al. Road traffic and childhood leukemia: the ESCALE study (SFCE). Environ Health Perspect 2011; 119: 566-72.

Amin R, Bohnert A, Holmes L, Rajasekaran A, Assanasen C. Epidemiologic mapping of Florida childhood cancer clusters. Pediatr Blood Cancer 2010; 54: 511-8.

Andersen CE, Raaschou-Nielsen O, Andersen HP, Lind M, Gravesen P, Thomsen BL, et al. Prediction of 222Rn in Danish dwellings using geology and house construction information from central databases. Radiat Prot Dosimetry 2007; 123: 83-94.

Appleton JD, Doyle E, Fenton D, Organo C. Radon potential mapping of the Tralee-Castleisland and Cavan areas (Ireland) based on airborne gamma-ray spectrometry and geology. J Radiol Prot 2011; 31: 221-35.

Appleton JD, Miles JCH. A statistical evaluation of the geogenic controls on indoor radon concentrations and radon risk. Journal of Environmental Radioactivity 2010; 101: 799-803.

Apte MG, Price PN, Nero AV, Revzan KL. Predicting New Hampshire indoor radon concentrations from geologic information and other covariates. Environmental Geology 1999; 37: 181-194.

Badaloni C, Ranucci A, Cesaroni G, Zanini G, Vienneau D, Al-Aidrous F, et al. Air pollution and childhood leukaemia: a nationwide case-control study in Italy. Occup Environ Med 2013; 70: 876-83.

Badr I, Oliver MA, Durrani SA. Statistical Evidence of the Geological Control over Radon Soil Gas Concentrations and its Implications for Mapping Radon Potential. Radiation Protection Dosimetry 1996; 63: 281-291.

Bailony MR, Hararah MK, Salhab AR, Ghannam I, Abdeen Z, Ghannam J. Cancer registration and healthcare access in West Bank, Palestine: a GIS analysis of childhood cancer, 1998-2007. Int J Cancer 2011; 129: 1180-9.

Bartley K, Metayer C, Selvin S, Ducore J, Buffler P. Diagnostic X-rays and risk of childhood leukaemia. Int J Epidemiol 2010; 39: 1628-37.

Baysson H, Billon S, Laurier D, Rogel A, Tirmarche M. Seasonal correction factors for estimating radon exposure in dwellings in France. Radiat Prot Dosimetry 2003; 104: 245-52.

Beck F, Richard JB, Deutsch A, Benmarhnia T, Pirard P, Roudier C, et al. Connaissance et perception du risque dû au radon en France. Cancer/Radiothérapie 2013; 17: 744-749.

BEIRVII. Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII ' Phase 2: National Academies Press, 2006a.

BEIRVII. Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII ' Phase 2: National Academies Press, 2006b.

Bellec S, Baccaini B, Goubin A, Rudant J, Ripert M, Hemon D, et al. Childhood leukaemia and population movements in France, 1990-2003. Br J Cancer 2008; 98: 225-31.

Bellec S, Hemon D, Rudant J, Goubin A, Clavel J. Spatial and space-time clustering of childhood acute leukaemia in France from 1990 to 2000: a nationwide study. Br J Cancer 2006; 94: 763-70.

Bernier MO, Mezzarobba M, Maupu E, Caer-Lorho S, Brisse HJ, Laurier D, et al. [Role of French hospital claims databases from care units in epidemiological studies: the example of the "Cohorte Enfant Scanner" study]. Rev Epidemiol Sante Publique 2012; 60: 363-70.

Bertolo A, Bigliotto C, Giovani C, Garavaglia M, Spinella M, Verdi L, et al. Spatial distribution of indoor radon in Triveneto (Northern Italy): a geostatistical approach. Radiation Protection Dosimetry 2009; 137: 318-323.

Bessou J, Deschamps F, Figueroa L, Cougnaud D. Methods used to estimate residential exposure to 50 Hz magnetic fields from overhead power lines in an epidemiological study in France. J Radiol Prot 2013; 33: 349-65.

Billon S, Morin A, Caer S, Baysson H, Gambard JP, Backe JC, et al. French population exposure to radon, terrestrial gamma and cosmic rays. Radiat Prot Dosimetry 2005; 113: 314-20.

Bithell JF, Dutton SJ, Draper GJ, Neary NM. Distribution of childhood leukaemias and non-Hodgkin's lymphomas near nuclear installations in England and Wales. BMJ 1994; 309: 501-5.

Bithell JF, Murphy MF, Stiller CA, Toumpakari E, Vincent T, Wakeford R. Leukaemia in young children in the vicinity of British nuclear power plants: a case-control study. Br J Cancer 2013; 109: 2880-5.

Bithell JF, Stewart AM. Pre-natal irradiation and childhood malignancy: a review of British data from the Oxford Survey. Br J Cancer 1975; 31: 271-87.

Black RJ, Sharp L, Harkness EF, McKinney PA. Leukaemia and non-Hodgkin's lymphoma: incidence in children and young adults resident in the Dounreay area of Caithness, Scotland in 1968-91. J Epidemiol Community Health 1994; 48: 232-6.

Boice JD, Jr., Miller RW. Childhood and adult cancer after intrauterine exposure to ionizing radiation. Teratology 1999; 59: 227-33.

Bonaventure A, Goujon-Bellec S, Rudant J, Orsi L, Leverger G, Baruchel A, et al. Maternal smoking during pregnancy, genetic polymorphisms of metabolic enzymes, and childhood acute leukemia: the ESCALE study (SFCE). Cancer Causes Control 2012; 23: 329-45.

Bonaventure A, Rudant J, Goujon-Bellec S, Orsi L, Leverger G, Baruchel A, et al. Childhood acute leukemia, maternal beverage intake during pregnancy, and metabolic polymorphisms. Cancer Causes Control 2013; 24: 783-93.

Borgoni R, Quatto P, Soma G, de Bartolo D. A geostatistical approach to define guidelines for radon prone area identification. Statistical Methods and Applications 2010; 19: 255-276.

Borgoni R, Tritto V, Bigliotto C, de Bartolo D. A geostatistical approach to assess the spatial association between indoor radon concentration, geological features and building characteristics: the case of Lombardy, Northern Italy. Int J Environ Res Public Health 2011; 8: 1420-40.

Bossew P, Dubois G, Tollefsen T. Investigations on indoor Radon in Austria, part 2: Geological classes as categorical external drift for spatial modelling of the Radon potential. Journal of Environmental Radioactivity 2008; 99: 81-97.

Bouges S, Daures JP, Hebrard M. [Incidence of acute leukemias, lymphomas and thyroid cancers in children under 15 years, living around the Marcoule nuclear site from 1985 to 1995]. Rev Epidemiol Sante Publique 1999; 47: 205-17.

Breslow NE, Day NE. Statistical methods in cancer research. Volume II--The design and analysis of cohort studies. IARC Sci Publ 1987: 1-406.

Brondum J, Shu XO, Steinbuch M, Severson RK, Potter JD, Robison LL. Parental cigarette smoking and the risk of acute leukemia in children. Cancer 1999; 85: 1380-8.

Brosselin P, Rudant J, Orsi L, Leverger G, Baruchel A, Bertrand Y, et al. Acute childhood leukaemia and residence next to petrol stations and automotive repair garages: the ESCALE study (SFCE). Occup Environ Med 2009; 66: 598-606.

Bunch KJ, Muirhead CR, Draper GJ, Hunter N, Kendall GM, O'Hagan JA, et al. Cancer in the offspring of female radiation workers: a record linkage study. Br J Cancer 2009; 100: 213-8.

Butland BK, Muirhead CR, Draper GJ. Radon and leukaemia. Lancet 1990; 335: 1338-1339.

Buttafuoco G, Tallarico A, Falcone G, Guagliardi I. A geostatistical approach for mapping and uncertainty assessment of geogenic radon gas in soil in an area of southern Italy. 2010.

Cardis E, Krewski D, Boniol M, Drozdovitch V, Darby SC, Gilbert ES, et al. Estimates of the cancer burden in Europe from radioactive fallout from the Chernobyl accident. Int J Cancer 2006; 119: 1224-35.

Carozza SE, Puumala SE, Chow EJ, Fox EE, Horel S, Johnson KJ, et al. Parental educational attainment as an indicator of socioeconomic status and risk of childhood cancers. Br J Cancer 2010; 103: 136-42.

Chan LC, Lam TH, Li CK, Lau YL, Yuen HL, Lee CW, et al. Is the timing of exposure to infection a major determinant of acute lymphoblastic leukaemia in Hong Kong? Paediatr Perinat Epidemiol 2002; 16: 154-65.

Cinelli G, Tondeur F, Dehandschutter B. Development of an indoor radon risk map of the Walloon region of Belgium, integrating geological information. Environmental Earth Sciences 2011; 62: 809-819.

Collman GW, Loomis DP, Sandler DP. Childhood cancer mortality and radon concentration in drinking water in North Carolina. Br J Cancer 1991; 63: 626-9.

Costa MA, Assunção RM, Kulldorff M. Constrained spanning tree algorithms for irregularly-shaped spatial clustering. Computational Statistics & Data Analysis 2012; 56: 1771-1783.

Crosignani P, Tittarelli A, Borgini A, Codazzi T, Rovelli A, Porro E, et al. Childhood leukemia and road traffic: A population-based case-control study. Int J Cancer 2004; 108: 596-9.

Darby SC, Whitley E, Howe GR, Hutchings SJ, Kusiak RA, Lubin JH, et al. Radon and cancers other than lung cancer in underground miners: a collaborative analysis of 11 studies. J Natl Cancer Inst 1995; 87: 378-84.

Davis S, Day RW, Kopecky KJ, Mahoney MC, McCarthy PL, Michalek AM, et al. Childhood leukaemia in Belarus, Russia, and Ukraine following the Chernobyl power station accident: results from an international collaborative population-based case-control study. Int J Epidemiol 2006; 35: 386-96.

Dockerty JD, Sharples KJ, Borman B. An assessment of spatial clustering of leukaemias and lymphomas among young people in New Zealand. J Epidemiol Community Health 1999a; 53: 154-8.

Dockerty JD, Skegg DC, Elwood JM, Herbison GP, Becroft DM, Lewis ME. Infections, vaccinations, and the risk of childhood leukaemia. Br J Cancer 1999b; 80: 1483-9.

Doll R, Wakeford R. Risk of childhood cancer from fetal irradiation. Br J Radiol 1997; 70: 130-9.

Drolet J-P, Martel R, Poulin P, Dessau J-C. Methodology developed to make the Quebec indoor radon potential map. Science of the Total Environment 2014; 473–474: 372-380.

Dubois G, Bossew P, Friedmann H. A geostatistical autopsy of the Austrian indoor radon survey (1992-2002). Sci Total Environ 2007; 377: 378-95.

Duczmal L, Cançado ALF, Takahashi RHC, Bessegato LF. A genetic algorithm for irregularly shaped spatial scan statistics. Computational Statistics & data Analysis 2007; 52: 43-52.

Elliott P, Toledano MB, Bennett J, Beale L, de Hoogh K, Best N, et al. Mobile phone base stations and early childhood cancers: case-control study. BMJ 2010; 340.

Evrard AS, Hemon D, Billon S, Laurier D, Jougla E, Tirmarche M, et al. Ecological association between indoor radon concentration and childhood leukaemia incidence in France, 1990-1998. European Journal of Cancer Prevention 2005; 14: 147-157.

Evrard AS, Hemon D, Billon S, Laurier D, Jougla E, Tirmarche M, et al. Childhood leukemia incidence and exposure to indoor radon, terrestrial and cosmic gamma radiation. Health Physics 2006a; 90: 569-579.

Evrard AS, Hemon D, Morin A, Laurier D, Tirmarche M, Backe JC, et al. Childhood leukaemia incidence around French nuclear installations using geographic zoning based on gaseous discharge dose estimates. Br J Cancer 2006b; 94: 1342-7.

Faure C, Mollie A, Bellec S, Guyot-Goubin A, Clavel J, Hemon D. Geographical variations in the incidence of childhood acute leukaemia in France over the period 1990-2004. Eur J Cancer Prev 2009; 18: 267-79.

Feychting M, Svensson D, Ahlbom A. Exposure to motor vehicle exhaust and childhood cancer. Scand J Work Environ Health 1998; 24: 8-11.

Foreman NK, Thorne R, Berry PJ, Oakhill A, Mott MG. Childhood malignancies in the south-west region of England, 1976-1985. Med Pediatr Oncol 1994; 23: 14-9.

Franco-Marina F, Villalba-Caloca J, Segovia N, Tavera L. Spatial indoor radon distribution in Mexico City. Science of the Total Environment 2003; 317: 91-103.

Friedmann H, Groller J. An approach to improve the Austrian Radon Potential Map by Bayesian statistics. Journal of Environmental Radioactivity 2010; 101: 804-808.

Gambard JP, Mitton N, Pirard P. Campagne nationale de mesure de l'exposition domestique au radon IPSN-DGS. Bilan et représentation cartographique des mesures au 01 janvier 2000. Institut de Protection et de Sûreté Nucléaire 2000.

Garcia-Talavera M, Garcia-Perez A, Rey C, Ramos L. Mapping radon-prone areas using gamma-radiation dose rate and geological information. J Radiol Prot 2013; 33: 605-20.

Gardner MJ, Snee MP, Hall AJ, Powell CA, Downes S, Terrell JD. Results of case-control study of leukaemia and lymphoma among young people near Sellafield nuclear plant in West Cumbria. BMJ 1990; 300: 423-9.

Gerken M, Kreienbrock L, Wellmann J, Kreuzer M, Wichmann HE. Models for retrospective quantification of indoor radon exposure in case-control studies. Health Phys 2000; 78: 268-78.

Gilbert ES. Ionising radiation and cancer risks: what have we learned from epidemiology? Int J Radiat Biol 2009; 85: 467-82.

Gilham C, Peto J, Simpson J, Roman E, Eden TO, Greaves MF, et al. Day care in infancy and risk of childhood acute lymphoblastic leukaemia: findings from UK case-control study. BMJ 2005; 330: 1294.

Goldsmith JR. Nuclear installations and childhood cancer in the UK: mortality and incidence for 0-9-year-old children, 1971-1980. Sci Total Environ 1992; 127: 13-35; discussion 43-55.

Goodman M, LaKind J, Fagliano J, Lash T, Wiemels J, Winn D, et al. Cancer Cluster Investigations: Review of the Past and Proposals for the Future. International Journal of Environmental Research and Public Health 2014; 11: 1479-1499.

Goujon-Bellec S, Demoury C, Guyot-Goubin A, Hemon D, Clavel J. Detection of clusters of a rare disease over a large territory: performance of cluster detection methods. Int J Health Geogr 2011; 10: 53.

Goujon-Bellec S, Mollie A, Rudant J, Guyot-Goubin A, Clavel J. Time trends and seasonal variations in the diagnosis of childhood acute lymphoblastic leukaemia in France. Cancer Epidemiol 2013; 37: 255-61.

Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer 2006; 6: 193-203.

Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. Leukemia 1988; 2: 120-5.

Greenland S, Sheppard AR, Kaune WT, Poole C, Kelsh MA. A pooled analysis of magnetic fields, wire codes, and childhood leukemia. Childhood Leukemia-EMF Study Group. Epidemiology 2000; 11: 624-34.

Gruber V, Bossew P, Cort MD, Tollefsen T. The European map of the geogenic radon potential. Journal of Radiological Protection 2013; 33: 51.

Gunby JA, Darby SC, Miles JCH, Green BMR, Cox DR. Factors affecting indoor radon concentrations in the United-Kingdom. Health Physics 1993; 64: 2-12.

Gundestrup M, Storm HH. Radiation-induced acute myeloid leukaemia and other cancers in commercial jet cockpit crew: a population-based cohort study. Lancet 1999; 354: 2029-31.

Harrison RM, Leung PL, Somervaille L, Smith R, Gilman E. Analysis of incidence of childhood cancer in the West Midlands of the United Kingdom in relation to proximity to main roads and petrol stations. Occup Environ Med 1999; 56: 774-80.

Hauri D, Spycher B, Huss A, Zimmermann F, Grotzer M, von der Weid N, et al. Domestic radon exposure and risk of childhood cancer: a prospective census-based cohort study. Environ Health Perspect 2013a; 121: 1239-44.

Hauri DD, Huss A, Zimmermann F, Kuehni CE, Roosli M. Prediction of residential radon exposure of the whole Swiss population: comparison of model-based predictions with measurement-based predictions. Indoor Air 2013b.

Hauri DD, Huss A, Zimmermann F, Kuehni CE, Röösli M. A prediction model for assessing residential radon concentration in Switzerland. Journal of Environmental Radioactivity 2012; 112: 83-89.

Hawkins MM, Wilson LM, Stovall MA, Marsden HB, Potok MH, Kingston JE, et al. Epipodophyllotoxins, alkylating agents, and radiation and risk of secondary leukaemia after childhood cancer. BMJ 1992; 304: 951-8.

Heck JE, Wu J, Lombardi C, Qiu J, Meyers TJ, Wilhelm M, et al. Childhood cancer and traffic-related air pollution exposure in pregnancy and early life. Environ Health Perspect 2013; 121: 1385-91.

Heinavaara S, Toikkanen S, Pasanen K, Verkasalo PK, Kurttio P, Auvinen A. Cancer incidence in the vicinity of Finnish nuclear power plants: an emphasis on childhood leukemia. Cancer Causes Control 2010; 21: 587-95.

Henshaw DL, Eatough JP, Richardson RB. Radon as a causative factor in induction of myeloid leukaemia and other cancers. Lancet 1990; 335: 1008-12.

Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N. Childhood leukaemia in Sweden: using GIS and a spatial scan statistic for cluster detection. Stat Med 1996; 15: 707-15.

Hoffmann W, Terschueren C, Richardson DB. Childhood leukemia in the vicinity of the Geesthacht nuclear establishments near Hamburg, Germany. Environ Health Perspect 2007; 115: 947-52.

Hsu WL, Preston DL, Soda M, Sugiyama H, Funamoto S, Kodama K, et al. The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950-2001. Radiat Res 2013; 179: 361-82.

Huang L, Pickle LW, Das B. Evaluating spatial methods for investigating global clustering and cluster detection of cancer cases. Stat Med 2008; 27: 5111-42.

Hunter N, Muirhead CR, Miles JCH, Appleton JD. Uncertainties in radon related to house-specific factors and proximity to geological boundaries in England. Radiation Protection Dosimetry 2009; 136: 17-22.

IARC. Some aromatic amines, anthraquinones and nitroso compounds, and inorganic fluorides used in drinking-water and dental preparations: International Agency for Research on Cancer, 1982.

IARC. Working Group on the Evaluation of Carcinogenic Risks to Humans. International Agency for Research on Cancer World Health Organization. Man-made mineral fibres and radon: International Agency for Research on Cancer, 1988.

IARC. Working Group on the Evaluation of Carcinogenic Risks to Humans International Agency for Research on Cancer World Health Organization. Ionizing radiation, part 1: x- and gamma (γ)-radiation, and neutrons: IARC, 2000.

IARC. Working Group on the Evaluation of Carcinogenic Risks to Humans. World Health Organization International Agency for Research on Cancer. Some Traditional Herbal Medicines, Some Mycotoxins, Naphthalene and Styrene: IARC Press, 2002.

ICRP. The 2007 Recommendations of the International Commission on Radiological Protection: Elsevier, published for the International Commission on Radiological Protection, 2008.

Ielsch G, Cushing ME, Combes P, Cuney M. Mapping of the geogenic radon potential in France to improve radon risk management: methodology and first application to region Bourgogne. J Environ Radioact 2010; 101: 813-20.

Ielsch G, Ferry C, Tymen G, Robé MC. Study of a predictive methodology for quantification and mapping of the radon-222 exhalation rate. Journal of Environmental Radioactivity 2002; 63: 15-33.

Ielsch G, Thiéblemont D, Labed V, Richon P, Tymen G, Ferry C, et al. Radon (222Rn) level variations on a regional scale: influence of the basement trace element (U, Th) geochemistry on radon exhalation rates. Journal of Environmental Radioactivity 2001; 53: 75-90.

Infante-Rivard C. Diagnostic x rays, DNA repair genes and childhood acute lymphoblastic leukemia. Health Phys 2003; 85: 60-4.

Infante-Rivard C, Fortier I, Olson E. Markers of infection, breast-feeding and childhood acute lymphoblastic leukaemia. Br J Cancer 2000; 83: 1559-64.

Infante-Rivard C, Krajinovic M, Labuda D, Sinnett D. Childhood acute lymphoblastic leukemia associated with parental alcohol consumption and polymorphisms of carcinogen-metabolizing genes. Epidemiology 2002; 13: 277-81.

Ivanov VK, Tsyb AF, Khait SE, Kashcheev VV, Chekin SY, Maksioutov MA, et al. Leukemia incidence in the Russian cohort of Chernobyl emergency workers. Radiat Environ Biophys 2012; 51: 143-9.

Johnson KJ, Alexander BH, Doody MM, Sigurdson AJ, Linet MS, Spector LG, et al. Childhood cancer in the offspring born in 1921-1984 to US radiologic technologists. Br J Cancer 2008; 99: 545-50.

Jourdan-Da Silva N, Perel Y, Mechinaud F, Plouvier E, Gandemer V, Lutz P, et al. Infectious diseases in the first year of life, perinatal characteristics and childhood acute leukaemia. Br J Cancer 2004; 90: 139-45.

Journy N, Ancelet S, Rehel JL, Mezzarobba M, Aubert B, Laurier D, et al. Predicted cancer risks induced by computed tomography examinations during childhood, by a quantitative risk assessment approach. Radiat Environ Biophys 2014; 53: 39-54.

Kaatsch P, Kaletsch U, Meinert R, Michaelis J. An extended study on childhood malignancies in the vicinity of German nuclear power plants. Cancer Causes Control 1998; 9: 529-33.

Kaatsch P, Spix C, Schulze-Rath R, Schmiedel S, Blettner M. Leukaemia in young children living in the vicinity of German nuclear power plants. Int J Cancer 2008; 122: 721-6.

Kaldor J. Second cancer following chemotherapy and radiotherapy. An epidemiological perspective. Acta Oncol 1990; 29: 647-55.

Kaletsch U, Kaatsch P, Meinert R, Schuz J, Czarwinski R, Michaelis J. Childhood cancer and residential radon exposure - results of a population-based case-control study in Lower Saxony (Germany). Radiation and Environmental Biophysics 1999; 38: 211-215.

Kemski J, Klingel R, Siehl A, Valdivia-Manchego M. From radon hazard to risk prediction-based on geological maps, soil gas and indoor measurements in Germany. Environmental Geology 2009; 56: 1269-1279.

Kemski J, Siehl A, Stegemann R, Valdivia-Manchego M. Mapping the geogenic radon potential in Germany. Science of the Total Environment 2001; 272: 217-230.

Kendall GM, Hughes JS, Oatway WB, Jones AL. Variations in radiation exposures of adults and children in the UK. J Radiol Prot 2006; 26: 257-76.

Kendall GM, Little MP, Wakeford R, Bunch KJ, Miles JC, Vincent TJ, et al. A record-based case-control study of natural background radiation and the incidence of childhood leukaemia and other cancers in Great Britain during 1980-2006. Leukemia 2012.

Kendall GM, Smith TJ. Doses to organs and tissues from radon and its decay products. J Radiol Prot 2002; 22: 389-406.

Kendall GM, Smith TJ. Doses from radon and its decay products to children. J Radiol Prot 2005; 25: 241-56.

Kesminiene A, Evrard AS, Ivanov VK, Malakhova IV, Kurtinaitis J, Stengrevics A, et al. Risk of hematological malignancies among Chernobyl liquidators. Radiat Res 2008; 170: 721-35.

Kinlen L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain. Lancet 1988; 2: 1323-7.

Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. Br J Cancer 1995; 71: 1-5.

Kinlen LJ. An examination, with a meta-analysis, of studies of childhood leukaemia in relation to population mixing. Br J Cancer 2012; 107: 1163-8.

Kinlen LJ, Bramald S. Paternal occupational contact level and childhood leukaemia in rural Scotland: a case-control study. Br J Cancer 2001; 84: 1002-7.

Kohli S, Noorlind Brage H, Lofman O. Childhood leukaemia in areas with different radon levels: a spatial and temporal analysis using GIS. J Epidemiol Community Health 2000; 54: 822-6.

Krestinina L, Preston DL, Davis FG, Epifanova S, Ostroumova E, Ron E, et al. Leukemia incidence among people exposed to chronic radiation from the contaminated Techa River, 1953-2005. Radiat Environ Biophys 2010; 49: 195-201.

Krestinina LY, Davis FG, Schonfeld S, Preston DL, Degteva M, Epifanova S, et al. Leukaemia incidence in the Techa River Cohort: 1953-2007. Br J Cancer 2013; 109: 2886-93.

Kroll ME, Stiller CA, Murphy MF, Carpenter LM. Childhood leukaemia and socioeconomic status in England and Wales 1976-2005: evidence of higher incidence in relatively affluent communities persists over time. Br J Cancer 2011; 105: 1783-7.

Kropat G, Bochud F, Jaboyedoff M, Laedermann J-P, Murith C, Palacios M, et al. Major influencing factors of indoor radon concentrations in Switzerland. Journal of Environmental Radioactivity 2014; 129: 7-22.

Kulldorff M, Huang L, Pickle L, Duczmal L. An elliptic spatial scan statistic. Stat Med 2006; 25: 3929-43.

Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. Stat Med 1995; 14: 799-810.

Kwan ML, Block G, Selvin S, Month S, Buffler PA. Food consumption by children and the risk of childhood acute leukemia. Am J Epidemiol 2004a; 160: 1098-107.

Kwan ML, Buffler PA, Abrams B, Kiley VA. Breastfeeding and the risk of childhood leukemia: a meta-analysis. Public Health Rep 2004b; 119: 521-35.

Lacour B, Guyot-Goubin A, Guissou S, Bellec S, Desandes E, Clavel J. Incidence of childhood cancer in France: National Children Cancer Registries, 2000-2004. Eur J Cancer Prev 2010; 19: 173-81.

Lalloue B, Monnez JM, Padilla C, Kihal W, Le Meur N, Zmirou-Navier D, et al. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. Int J Equity Health 2013; 12: 21.

Laurent O, Ancelet S, Richardson DB, Hemon D, Ielsch G, Demoury C, et al. Potential impacts of radon, terrestrial gamma and cosmic rays on childhood leukemia in France: a quantitative risk assessment. Radiat Environ Biophys 2013; 52: 195-209.

Laurier D, Bard D. Epidemiologic studies of leukemia among persons under 25 years of age living near nuclear sites. Epidemiol Rev 1999; 21: 188-206.

Laurier D, Jacob S, Bernier MO, Leuraud K, Metz C, Samson E, et al. Epidemiological studies of leukaemia in children and young adults around nuclear facilities: a critical review. Radiat Prot Dosimetry 2008; 132: 182-90.

Laurier D, Valenty M, Tirmarche M. Radon exposure and the risk of leukemia: a review of epidemiological studies. Health Phys 2001; 81: 272-88.

Lehmann R, Kemski J, Siehl A, Stegemann R, Valdivia-Manchego M. The regional distribution of indoor radon concentration in Germany. International Congress Series 2002; 1225: 55-61.

Little MP. Cancer after exposure to radiation in the course of treatment for benign and malignant disease. Lancet Oncol 2001a; 2: 212-20.

Little MP. Comparison of the risks of cancer incidence and mortality following radiation therapy for benign and malignant disease with the cancer risks observed in the Japanese A-bomb survivors. Int J Radiat Biol 2001b; 77: 431-64.

Little MP. Leukaemia following childhood radiation exposure in the Japanese atomic bomb survivors and in medically exposed groups. Radiat Prot Dosimetry 2008; 132: 156-65.

Little MP, Wakeford R, Kendall GM. Updated estimates of the proportion of childhood leukaemia incidence in Great Britain that may be caused by natural background ionising radiation. J Radiol Prot 2009; 29: 467-82.

Little MP, Wakeford R, Lubin JH, Kendall GM. The statistical power of epidemiological studies analyzing the relationship between exposure to ionizing radiation and cancer, with special reference to childhood leukemia and natural background radiation. Radiat Res 2010; 174: 387-402.

Lombardi C, Heck JE, Cockburn M, Ritz B. Solar UV radiation and cancer in young children. Cancer Epidemiol Biomarkers Prev 2013; 22: 1118-28.

Lowengart RA, Peters JM, Cicioni C, Buckley J, Bernstein L, Preston-Martin S, et al. Childhood leukemia and parents' occupational and home exposures. J Natl Cancer Inst 1987; 79: 39-46.

Lubin JH, Linet MS, Boie JD, Buckley J, Conrath SM, Hatch EE, et al. Case-control study of childhood acute lymphoblastic leukemia and residential radon exposure. Journal of the National Cancer Institute 1998; 90: 294-300.

Lucie NP. Radon exposure and leukemia. Lancet 1989; 2: 99-100.

Ma X, Buffler PA, Gunier RB, Dahl G, Smith MT, Reinier K, et al. Critical windows of exposure to household pesticides and risk of childhood leukemia. Environ Health Perspect 2002; 110: 955-60.

Maged AF, Mokhtar GM, El-Tobgui MM, Gabbr AA, Attia NI, Abu Shady MM. Domestic radon concentration and childhood cancer study in Cairo, Egypt. Environmental Carcinogenesis & Ecotoxicology Reviews-Part C of Journal of Environmental Science and Health 2000; 18: 153-170.

Makelainen I, Arvela H, Voutilainen A. Correlations between radon concentration and indoor gamma dose rate, soil permeability and dwelling substructure and ventilation. Science of the Total Environment 2001; 272: 283-289.

Mäkeläinen I, Kinnunen T, Reisbacka H, Valmari T, Arvela H. Radon in Finnish dwellings - Sample survey 2006. Radiation and Nuclear Safety Authority, Helsinki, 2009.

Marsh JW, Bessa Y, Birchall A, Blanchardon E, Hofmann W, Nosske D, et al. Dosimetric models used in the Alpha-Risk project to quantify exposure of uranium miners to radon gas and its progeny. Radiat Prot Dosimetry 2008; 130: 101-6.

Mathews JD, Forsythe AV, Brady Z, Butler MW, Goergen SK, Byrnes GB, et al. Cancer risk in 680,000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. BMJ 2013; 346: f2360.

McNally RJ, Alexander FE, Bithell JF. Space-time clustering of childhood cancer in great Britain: a national study, 1969-1993. Int J Cancer 2006; 118: 2840-6.

McNally RJ, Alexander FE, Vincent TJ, Murphy MF. Spatial clustering of childhood cancer in Great Britain during the period 1969-1993. Int J Cancer 2009; 124: 932-6.

McNally RJ, Eden TO. An infectious aetiology for childhood acute leukaemia: a review of the evidence. Br J Haematol 2004; 127: 243-63.

Mechenet V, Debayle C. Modernisation du réseau de dosimétrie environnementale. Institut de Radioprotection et de Sûreté Nucléaire, 2013.

Menegaux F, Baruchel A, Bertrand Y, Lescoeur B, Leverger G, Nelken B, et al. Household exposure to pesticides and risk of childhood acute leukaemia. Occup Environ Med 2006; 63: 131-4.

Menegaux F, Steffen C, Bellec S, Baruchel A, Lescoeur B, Leverger G, et al. Maternal coffee and alcohol consumption during pregnancy, parental smoking and risk of childhood acute leukaemia. Cancer Detect Prev 2005; 29: 487-93.

Miles JCH, Appleton JD. Mapping variation in radon potential both between and within geological units. Journal of Radiological Protection 2005; 25: 257-276.

Minda M, Toth G, Horvath I, Barnet I, Hamori K, Toth E. Indoor radon mapping and its relation to geology in Hungary. Environmental Geology 2009; 57: 601-609.

Mohner M, Gellissen J, Marsh JW, Gregoratto D. Occupational and diagnostic exposure to ionizing radiation and leukemia risk among German uranium miners. Health Phys 2010; 99: 314-21.

Mohner M, Lindtner M, Otten H, Gille HG. Leukemia and exposure to ionizing radiation among German uranium miners. Am J Ind Med 2006; 49: 238-48.

Mora P, Picado E, Minato S. Natural radiation doses for cosmic and terrestrial components in Costa Rica. Applied Radiation and Isotopes 2007; 65: 79-84.

Moran PA. Rank correlation and product-moment correlation. Biometrika 1948; 35: 203-6.

Mosavi-Jarrahi A, Moini M, Mohagheghi MA, Alebouyeh M, Yazdizadeh B, Shahabian A, et al. Clustering of childhood cancer in the inner city of Tehran metropolitan area: a GIS-based analysis. Int J Hyg Environ Health 2007; 210: 113-9.

Muirhead CR, Butland BK. Methods for investigating localized clustering of disease. Testing for over-dispersion using an adapted form of the Potthoff-Whittinghill method. IARC Sci Publ 1996: 40-52; discussion 155-63.

Muirhead CR, Butland BK, Green BM, Draper G. Childhood leukaemia and natural radiation. Lancet 1991; 337: 503-504.

Musselman JR, Spector LG. Childhood cancer incidence in relation to sunlight exposure. Br J Cancer 2011; 104: 214-20.

NCRP. Ionizing Radiation Exposure of the Population of the United States: Recommendations of the National Council on Radiation Protection and Measurements, March 3, 2009: National Council on Radiation Protection and Measurements, 2009.

Nordlinder R, Jarvholm B. Environmental exposure to gasoline and leukemia in children and young adults--an ecology study. Int Arch Occup Environ Health 1997; 70: 57-60.

Noshchenko AG, Bondar OY, Drozdova VD. Radiation-induced leukemia among children aged 0-5 years at the time of the Chernobyl accident. Int J Cancer 2010; 127: 412-26.

Noshchenko AG, Zamostyan PV, Bondar OY, Drozdova VD. Radiation-induced leukemia risk among those aged 0-20 at the time of the Chernobyl accident: a case-control study in the Ukraine. Int J Cancer 2002; 99: 609-18.

Nyari TA, Ottoffy G, Bartyik K, Thurzo L, Solymosi N, Cserni G, et al. Spatial clustering of childhood acute lymphoblastic leukaemia in hungary. Pathol Oncol Res 2013; 19: 297-302.

Orsi L, Rudant J, Bonaventure A, Goujon-Bellec S, Corda E, Evans TJ, et al. Genetic polymorphisms and childhood acute lymphoblastic leukemia: GWAS of the ESCALE study (SFCE). Leukemia 2012; 26: 2561-4.

Ozasa K, Shimizu Y, Suyama A, Kasagi F, Soda M, Grant EJ, et al. Studies of the mortality of atomic bomb survivors, Report 14, 1950-2003: an overview of cancer and noncancer diseases. Radiat Res 2012; 177: 229-43.

Papaefthymiou H, Mavroudis A, Kritidis P. Indoor radon levels and influencing factors in houses of Patras, Greece. Journal of Environmental Radioactivity 2003; 66: 247-260.

Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 2009; 41: 1006-10.

Pearce MS, Salotti JA, Little MP, McHugh K, Lee C, Kim KP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. Lancet 2012; 380: 499-505.

Peh SC, Nadarajah VS, Tai YC, Kim LH, Abdullah WA. Pattern of Epstein-Barr virus association in childhood non-Hodgkin's lymphoma: experience of university of malaya medical center. Pathol Int 2004; 54: 151-7.

Perrillat F, Clavel J, Auclerc MF, Baruchel A, Leverger G, Nelken B, et al. Day-care, early common infections and childhood acute leukaemia: a multicentre French case-control study. Br J Cancer 2002; 86: 1064-9.

Peters JM, Preston-Martin S, London SJ, Bowman JD, Buckley JD, Thomas DC. Processed meats and risk of childhood leukemia (California, USA). Cancer Causes Control 1994; 5: 195-202.

Petoussi N, Jacob P, Zankl M, Saito K. Organ Doses for Foetuses, Babies, Children and Adults from Environmental Gamma Rays. Radiation Protection Dosimetry 1991; 37: 31-41.

Petridou E, Alexander FE, Trichopoulos D, Revinthi K, Dessypris N, Wray N, et al. Aggregation of childhood leukemia in geographic areas of Greece. Cancer Causes Control 1997; 8: 239-45.

Petridou E, Dalamaga M, Mentis A, Skalkidou A, Moustaki M, Karpathios T, et al. Evidence on the infectious etiology of childhood leukemia: the role of low herd immunity (Greece). Cancer Causes Control 2001; 12: 645-52.

Pirard P, Richardson S, Laurier D, Mitton N, Gambard J, Monfort C, et al. Analyse écologique de la relation entre l'exposition au radon dans l'habitat et la mortalité par cancer du poumon en France. Rapport final dans le cadre du Programme Environnement-Santé 1996, Convention EN96C5 et EN96D5. Institut de Protection et de Sûreté Nucléaire, SEGR/LEADS/NT/2000-23, 2000.

Poje M, Vuković B, Radolić V, Miklavčić I, Faj D, Varga Pajtler M, et al. Mapping of cosmic radiation dose in Croatia. Journal of Environmental Radioactivity 2012; 103: 30-33.

Poole C, Greenland S, Luetters C, Kelsey JL, Mezei G. Socioeconomic status and childhood leukaemia: a review. Int J Epidemiol 2006; 35: 370-84.

Pornet C, Delpierre C, Dejardin O, Grosclaude P, Launay L, Guittet L, et al. Construction of an adaptable European transnational ecological deprivation index: the French version. J Epidemiol Community Health 2012; 66: 982-9.

Potthoff RF, Whittinghill M. Testing for homogeneity. II. The Poisson distribution. Biometrika 1966; 53: 183-90.

Pouchol C, Ielsch G. Analyse spatiale et cartographie du radon sur le territoire métropolitain par l'utilisation de méthodes géostatistiques. Institut de Radioprotection et de Sûreté Nucléaire, 2012.

Preston DL, Kusumi S, Tomonaga M, Izumi S, Ron E, Kuramoto A, et al. Cancer incidence in atomic bomb survivors. Part III. Leukemia, lymphoma and multiple myeloma, 1950-1987. Radiat Res 1994; 137: S68-97.

Pukkala E, Aspholm R, Auvinen A, Eliasch H, Gundestrup M, Haldorsen T, et al. Cancer incidence among 10,211 airline pilots: a Nordic study. Aviat Space Environ Med 2003; 74: 699-706.

Quindos Poncela LS, Fernandez PL, Gomez Arozamena J, Sainz C, Fernandez JA, Suarez Mahou E, et al. Natural gamma radiation map (MARNA) and indoor radon levels in Spain. Environ Int 2004; 29: 1091-6.

Raaschou-Nielsen O. Indoor radon and childhood leukaemia. Radiat Prot Dosimetry 2008; 132: 175-81.

Raaschou-Nielsen O, Andersen CE, Andersen HP, Gravesen P, Lind M, Schuz J, et al. Domestic radon and childhood cancer in Denmark. Epidemiology 2008; 19: 536-43.

Raaschou-Nielsen O, Hertel O, Thomsen BL, Olsen JH. Air pollution from traffic at the residence of children with cancer. Am J Epidemiol 2001; 153: 433-43.

Raaschou-Nielsen O, Reynolds P. Air pollution and childhood cancer: a review of the epidemiological literature. Int J Cancer 2006; 118: 2920-9.

Rajaraman P, Simpson J, Neta G, Berrington de Gonzalez A, Ansell P, Linet MS, et al. Early life exposure to diagnostic radiation and ultrasound scans and risk of childhood cancer: case-control study. BMJ 2011; 342: d472.

Rannou A, Aubert B, Scanff P. Exposition de la population française aux rayonnements ionisants. Institut de Radioprotection et de Sûreté Nucléaire, 2006.

Rannou A, Madelmont C, Renouard H. Survey of natural radiation in France. Sci Total Environ 1985; 45: 467-474.

Raspa G, Salvi F, Torri G. Probability mapping of indoor radon-prone areas unsing disjunctive kriging. Radiation Protection Dosimetry 2010; 138: 3-19.

Rericha V, Kulich M, Rericha R, Shore DL, Sandler DP. Incidence of leukemia, lymphoma, and multiple myeloma in Czech uranium miners: a case-cohort study. Environ Health Perspect 2006; 114: 818-22.

Rey G, Jougla E, Fouillet A, Hemon D. Ecological association between a deprivation index and mortality in France over the period 1997 - 2001: variations with spatial scale, degree of urbanicity, age, gender and cause of death. BMC Public Health 2009; 9: 33.

Reynolds P, Von Behren J, Gunier RB, Goldberg DE, Hertz A. Residential exposure to traffic in California and childhood cancer. Epidemiology 2004; 15: 6-12.

Reynolds P, Von Behren J, Gunier RB, Goldberg DE, Hertz A, Smith D. Traffic patterns and childhood cancer incidence rates in California, United States. Cancer Causes Control 2002; 13: 665-73.

Richardson D, Sugiyama H, Nishi N, Sakata R, Shimizu Y, Grant EJ, et al. Ionizing radiation and leukemia mortality among Japanese Atomic Bomb Survivors, 1950-2000. Radiat Res 2009; 172: 368-82.

Richardson S, Monfort C, Green M, Draper G, Muirhead C. Spatial variation of natural radiation and childhood leukaemia incidence in Great Britain. Stat Med 1995; 14: 2487-501.

Rogerson PA. The Detection of Clusters Using a Spatial Version of the Chi-Square Goodness-of-Fit Statistic. Geographical Analysis 1999; 31: 130-147.

Rommens C, Ringeard C, Hubert P. Exposure of red bone marrow to ionising radiation from natural and medical sources in France. J Radiol Prot 2001; 21: 209-19.

Ron E. Cancer risks from medical radiation. Health Phys 2003; 85: 47-59.

Ron E, Modan B, Boice JD, Jr. Mortality after radiotherapy for ringworm of the scalp. Am J Epidemiol 1988; 127: 713-25.

Ross JA, Potter JD, Reaman GH, Pendergrass TW, Robison LL. Maternal exposure to potential inhibitors of DNA topoisomerase II and infant leukemia (United States): a report from the Children's Cancer Group. Cancer Causes Control 1996; 7: 581-90.

Rudant J, Baccaini B, Ripert M, Goubin A, Bellec S, Hemon D, et al. Population-mixing at the place of residence at the time of birth and incidence of childhood leukaemia in France. Eur J Cancer 2006; 42: 927-33.

Rudant J, Menegaux F, Leverger G, Baruchel A, Lambilliotte A, Bertrand Y, et al. Childhood hematopoietic malignancies and parental use of tobacco and alcohol: the ESCALE study (SFCE). Cancer Causes Control 2008; 19: 1277-90.

Rudant J, Menegaux F, Leverger G, Baruchel A, Nelken B, Bertrand Y, et al. Family history of cancer in children with acute leukemia, Hodgkin's lymphoma or non-Hodgkin's lymphoma: the ESCALE study (SFCE). Int J Cancer 2007a; 121: 119-26.

Rudant J, Menegaux F, Leverger G, Baruchel A, Nelken B, Bertrand Y, et al. Household exposure to pesticides and risk of childhood hematopoietic malignancies: The ESCALE study (SFCE). Environ Health Perspect 2007b; 115: 1787-93.

Rybach L, Bachler D, Bucher B, Schwarz G. Radiation doses of Swiss population from external sources. J Environ Radioact 2002; 62: 277-86.

Savitz DA, Feingold L. Association of childhood cancer with residential traffic density. Scand J Work Environ Health 1989; 15: 360-3.

Schmiedel S, Blettner M, Kaatsch P, Schuz J. Spatial clustering and space-time clusters of leukemia among children in Germany, 1987-2007. Eur J Epidemiol 2010; 25: 627-33.

Schmiedel S, Jacquez GM, Blettner M, Schuz J. Spatial clustering of leukemia and type 1 diabetes in children in Denmark. Cancer Causes Control 2011.

Schonfeld SJ, Tsareva YV, Preston DL, Okatenko PV, Gilbert ES, Ron E, et al. Cancer mortality following in utero exposure among offspring of female Mayak Worker Cohort members. Radiat Res 2012; 178: 160-5.

Schraube H, Leuthold G, Roesler S, Heinrich W. Neutron spectra at flight altitudes and their radiological estimation. Adv Space Res 1998; 21: 1727-38.

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.

Sermage-Faure C, Laurier D, Goujon-Bellec S, Chartier M, Guyot-Goubin A, Rudant J, et al. Childhood leukemia around French nuclear power plants--the Geocap study, 2002-2007. Int J Cancer 2012; 131: E769-80.

Severson RK, Buckley JD, Woods WG, Benjamin D, Robison LL. Cigarette smoking and alcohol consumption by parents of children with acute myeloid leukemia: an analysis within morphological subgroups--a report from the Childrens Cancer Group. Cancer Epidemiol Biomarkers Prev 1993; 2: 433-9.

Shahbazi-Gahrouei D, Gholami M, Setayandeh S. A review on natural background radiation. Adv Biomed Res 2013; 2: 65.

Shi X, Hoftiezer DJ, Duell EJ, Onega TL. Spatial association between residential radon concentration and bedrock types in New Hampshire. Environmental Geology 2006; 51: 65-71.

Shilnikova NS, Preston DL, Ron E, Gilbert ES, Vassilenko EK, Romanov SA, et al. Cancer mortality risk among workers at the Mayak nuclear complex. Radiat Res 2003; 159: 787-98.

Shore RE, Moseson M, Harley N, Pasternack BS. Tumors and other diseases following childhood x-ray treatment for ringworm of the scalp (Tinea capitis). Health Phys 2003; 85: 404-8.

Shu XO, Ross JA, Pendergrass TW, Reaman GH, Lampkin B, Robison LL. Parental alcohol consumption, cigarette smoking, and risk of infant leukemia: a Childrens Cancer Group study. J Natl Cancer Inst 1996; 88: 24-31.

Smith BJ, Zhang L, Field RW. Iowa radon leukaemia study: a hierarchical population risk model for spatially correlated exposure measured with error. Stat Med 2007; 26: 4619-42.

Smith M. Considerations on a possible viral etiology for B-precursor acute lymphoblastic leukemia of childhood. J Immunother 1997; 20: 89-100.

Smith MA, Simon R, Strickler HD, McQuillan G, Ries LA, Linet MS. Evidence that childhood acute lymphoblastic leukemia is associated with an infectious agent linked to hygiene conditions. Cancer Causes Control 1998; 9: 285-98.

Song C, Kulldorff M. Power evaluation of disease clustering tests. Int J Health Geogr 2003; 2: 9.

Spix C, Schmiedel S, Kaatsch P, Schulze-Rath R, Blettner M. Case-control study on childhood cancer in the vicinity of nuclear power plants in Germany 1980-2003. Eur J Cancer 2008; 44: 275-84.

Spycher BD, Feller M, Zwahlen M, Roosli M, von der Weid NX, Hengartner H, et al. Childhood cancer and nuclear power plants in Switzerland: a census-based cohort study. Int J Epidemiol 2011; 40: 1247-60.

Steenland K, Deddens JA, Zhao S. Biases in estimating the effect of cumulative exposure in log-linear models when estimated exposure levels are assigned. Scand J Work Environ Health 2000; 26: 37-43.

Steffen C, Auclerc MF, Auvrignon A, Baruchel A, Kebaili K, Lambilliotte A, et al. Acute childhood leukaemia and environmental exposure to potential sources of benzene and other hydrocarbons; a case-control study. Occup Environ Med 2004; 61: 773-8.

Steinbuch M, Weinberg CR, Buckley JD, Robison LL, Sandler DP. Indoor residential radon exposure and risk of childhood acute myeloid leukaemia. Br J Cancer 1999; 81: 900-6.

Stiller CA. Epidemiology and genetics of childhood cancer. Oncogene 2004; 23: 6429-44.

Stjernfeldt M, Samuelsson L, Ludvigsson J. Radiation in dwellings and cancer in children. Pediatr Hematol Oncol 1987; 4: 55-61.

Sundal AV, Henriksen H, Soldal O, Strand T. The influence of geological factors on indoor radon concentrations in Norway. Science of the Total Environment 2004; 328: 41-53.

Szabo KZ, Jordan G, Horvath A, Szabo C. Mapping the geogenic radon potential: methodology and spatial analysis for central Hungary. J Environ Radioact 2014; 129: 107-20.

Tango T. A spatial scan statistic with a restricted likelihood ratio. Japanese Journal of Biometrics 2008; 29: 75-95.

Tango T, Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. Int J Health Geogr 2005; 4: 11.

Thorne R, Foreman NK, Mott MG. Radon in Devon and Cornwall and paediatric malignancies. European Journal of Cancer 1996; 32: 282-285.

Timonen T, Nayha S, Koskela T, Pukkala E. Are sunlight deprivation and influenza epidemics associated with the onset of acute leukemia? Haematologica 2007; 92: 1553-6.

Tong J, Qin LQ, Cao Y, Li JX, Zhang J, Nie JH, et al. Environmental radon exposure and childhood leukemia. Journal of Toxicology and Environmental Health-Part B-Critical Reviews 2012; 15: 332-347.

Tucker MA, Fraumeni JF, Jr. Treatment-related cancers after gynecologic malignancy. Cancer 1987; 60: 2117-22.

Turner MC, Wigle DT, Krewski D. Residential pesticides and childhood leukemia: a systematic review and meta-analysis. Environ Health Perspect 2010; 118: 33-41.

UKCCS. The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: 2: gamma radiation. British Journal of Cancer 2002a; 86: 1727-1731.

UKCCS. The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: I: radon gas. British Journal of Cancer 2002b; 86: 1721-1726.

UNSCEAR. Sources and Effects of Ionizing Radiation: UNSCEAR 2000 Report to the General Assembly, with Scientific Annex: Bernan Assoc, 2000.

UNSCEAR. Effects of Ionizing Radiation: United Nations Scientific Committee on the Effects of Atomic Radiation - UNSCEAR 2006 Report, Volume II - Report to the General Assembly, with Scientific Annexes C, D, and E: UN, 2006.

UNSCEAR. Sources and Effects of Ionizing Radiation, Unscear 2008 Report: United Nations Publications, 2008.

Valmari T, Makelainen I, Reisbacka H, Arvela H. Finnish radon situation analysed using national measurement database. Radiation Protection Dosimetry 2011; 145: 101-106.

Van Maele-Fabry G, Lantin AC, Hoet P, Lison D. Residential exposure to pesticides and childhood leukaemia: a systematic review and meta-analysis. Environ Int 2011; 37: 280-91.

Verger P C-CL. Evaluation des conséquences sanitaires de l'accident de Tchernobyl en France: dispositif de surveillance épidémiologique, état des connaissances, évaluation des risques et perspectives. IPSN, InVS, 2009, pp. 22.

Vinson F, Merhi M, Baldi I, Raynal H, Gamet-Payrastre L. Exposure to pesticides and risk of childhood cancer: a meta-analysis of recent epidemiological studies. Occup Environ Med 2011; 68: 694-702.

Visser O, van Wijnen JH, van Leeuwen FE. Residential traffic density and cancer incidence in Amsterdam, 1989-1997. Cancer Causes Control 2004; 15: 331-9.

Von Behren J, Reynolds P, Gunier RB, Rull RP, Hertz A, Urayama KY, et al. Residential traffic density and childhood leukemia risk. Cancer Epidemiol Biomarkers Prev 2008; 17: 2298-301.

Wackernagel H. Multivariate geostatistics: Springer, 2003.

Wakeford R. Childhood leukaemia following medical diagnostic exposure to ionizing radiation in utero or after birth. Radiat Prot Dosimetry 2008; 132: 166-74.

Wakeford R. The risk of childhood leukaemia following exposure to ionising radiation--a review. J Radiol Prot 2013; 33: 1-25.

Wakeford R, Kendall GM, Little MP. The proportion of childhood leukaemia incidence in Great Britain that may be caused by natural background ionizing radiation. Leukemia 2009; 23: 770-6.

Wakeford R, Little MP. Risk coefficients for childhood cancer after intrauterine irradiation: a review. Int J Radiat Biol 2003; 79: 293-309.

Wakeford R, Little MP, Kendall GM. Risk of childhood leukemia after low-level exposure to ionizing radiation. Expert Rev Hematol 2010; 3: 251-4.

Waller LA, Hill EG, Rudd RA. The geography of power: statistical performance of tests of clusters and clustering in heterogeneous populations. Stat Med 2006; 25: 853-65.

Wheeler DC. A comparison of spatial clustering and cluster detection techniques for childhood leukemia incidence in Ohio, 1996-2003. Int J Health Geogr 2007; 6: 13.

White-Koning ML, Hemon D, Laurier D, Tirmarche M, Jougla E, Goubin A, et al. Incidence of childhood leukaemia in the vicinity of nuclear sites in France, 1990-1998. Br J Cancer 2004; 91: 916-22.

Wigle DT, Turner MC, Krewski D. A systematic review and meta-analysis of childhood leukemia and parental occupational pesticide exposure. Environ Health Perspect 2009; 117: 1505-13.

Wrixon AD, Board GBNRP. Natural Radiation Exposure in UK Dwellings: National Radiological Protection Board, 1988.

Yoshimoto Y, Neel JV, Schull WJ, Kato H, Soda M, Eto R, et al. Malignant tumors during the first 2 decades of life in the offspring of atomic bomb survivors. Am J Hum Genet 1990; 46: 1041-52.

Youlden DR, Baade PD, Valery PC, Hassall TE, Ward LJ, Green AC, et al. Area-based differentials in childhood cancer incidence in Australia, 1996-2006. Pediatr Blood Cancer 2012; 58: 390-4.

Zablotska LB, Lane RS, Frost SE. Mortality (1950-1999) and cancer incidence (1969-1999) of workers in the Port Hope cohort study exposed to a unique combination of radium, uranium and gamma-ray doses. BMJ Open 2013; 3.

Zahm SH, Ward MH. Pesticides and childhood cancer. Environ Health Perspect 1998; 106 Suppl 3: 893-908.

Zeeb H, Hammer GP, Blettner M. Epidemiological investigations of aircrew: an occupational group with low-level cosmic radiation exposure. J Radiol Prot 2012; 32: N15-9.

Zhu HC, Charlet JM, Poffijn A. Radon risk mapping in southern Belgium: an application of geostatistical and GIS techniques. Science of the Total Environment 2001; 272: 203-210.

Zhukovsky M, Onishchenko A, Varaksin A, Vasilyev A. The influence of radon measurment errors on the uncertainties of epidemiological case-control studies. Radiation Protection Dosimetry 2011; 145: 243-247.

Zunic ZS, Yarmoshenko IV, Birovljev A, Bochicchio F, Quarto M, Obryk B, et al. Radon survey in the high natural radiation region of Niska Banja, Serbia. J Environ Radioact 2007; 92: 165-74.

Partie IX Annexes

Tableau 21. Principales études rapportant un taux d'incidence chez les enfants exposés aux radiations (UNSCEAR 2006)

Etude			ERR [a] moyen pour 1 Sv	EAR moyen (10 000 PA Sv)$^{-1}$
Exposition externe				
-LSS (Preston et al., 1994)				
	sexe	H	4,66 [3,07 ; 6,88]	4,14 [3,06 ; 5,39]
		F	5,05 [3,24 ; 7,61]	2,41 [1,71 ; 3,23]
	âge à l'exposition	< 20 ans	8,27 [4,95 ;13,66]	2,79 [1,99 ; 3,74]
		20 – 39 ans	3,59 [2,01 ; 5,97]	2,69 [1,70 ; 3,90]
		≥ 40 ans	3,98 [2,32 ; 6,45]	4,68 [3,10 ; 6,57]
-LSS (Hsu et al., 2013)				
	âge à l'exposition	< 20 ans	6,5 [4,0 ; 10,3] [b]	-
		20 – 39 ans	3,9 [2,3 ; 6,1] [b]	-
		≥ 40 ans	4,0 [2,1 ; 6,9] [b]	-
-Scanner pédiatrique (Pearce et al., 2012) [b]			36 [5 ; 120] [d]	-
-Scanner pédiatrique (Mathews et al., 2013)			39 [14 ; 70] [d]	-
-Hodgkin's disease (Kaldor et al., 1990) [c]			0,24 [0,04 ; 0,43] [b]	-
-UKCC (Hawkins et al., 1992) [e]			0,24 [0,01 ; 1,28] [b]	-
-International Childhood cancer (Tucker et al., 1987)			0,0 [0,00 ; 0,004]	-
Exposition interne				
-Tchernobyl (Biélorussie, Russie, Ukraine) (Davis et al., 2006)			32,4 [9,78 ; 84,0]	-
-Tchernobyl (Ukraine) (Noshchenko et al., 2002)			2,5 [1 ; 5,4] [b,f,g]	-
-Tchernobyl (Ukraine) (Noshchenko et al., 2010)			2,1 [1,2 ; 3,7] [b,d,f]	-
			4,4 [1,3 ; 15]	-

[a] Intervalle de confiance à 90%
[b] Intervalle de confiance à 95%
[c] Estimation de risque basée sur (Little, 2001b)
[d] ERR par Gy pour leucémie et myelodysplasie
[e] estimation de risque basée sur l'UNSCEAR 2000
[f] OR
[g] RR parmi chez les personnes ayant reçu plus de 10 mSv ou plus par rapport à celles ayant reçu moins de 2 mSv

Partie X Articles

Title:
Spatial variations of childhood acute leukaemia in France, 1990-2006: global spatial heterogeneity and cluster detection at 'living zone' level

Short title:
Spatial heterogeneity in childhood leukaemia

Authors:
Claire Demoury[a, b], Stéphanie Goujon-Bellec[a, b, c], Aurélie Guyot-Goubin[a, b, c], Denis Hémon[a, b] and Jacqueline Clavel[a, b, c]

[a] INSERM, CESP Environmental Epidemiology of Cancer, U1018, 94807 Villejuif, France

[b] University Paris-Sud 11, UMRS-1018, 94807 Villejuif, France

[c] French National Registry of Childhood Haematopoietic malignancies (NRCH), 94807 Villejuif, France

Correspondence to Jacqueline Clavel,
Centre de recherche en Epidémiologie et Santé des Populations (CESP),
Epidémiologie environnementale des cancers UMRS-1018, Equipe 6 INSERM,
16, Avenue Paul Vaillant-Couturier,
F-94807 Villejuif Cedex, France
Tel: +33 1 45 59 50 37; fax: +33 1 45 59 51 51
e-mail: jacqueline.clavel@inserm.fr

Conflicts of interest: None declared

Author's copy

Abstract

Objectives Childhood acute leukaemia (AL) accounts for a third of childhood cancers. Analysing the spatial distribution of the incidence of AL and its lymphoblastic and myeloblastic subtypes may contribute to the identification of risk factors. This national registry-based study aimed to evaluate global spatial heterogeneity in the incidence rates of AL and to detect clusters in France over the period 1990-2006 on the 'living zone' (LZ) scale.

Methods Between 1990 and 2006, 7,675 cases of AL were registered in the National Registry of Childhood Haematopoietic malignancies. Their spatial distribution in the 1,895 LZ was first evaluated with two tests for global spatial heterogeneity (Potthoff-Witthinghill and Rogerson's tests) and then with the SaTScan and FleXScan methods, which aim to locate spatial and space-time clusters.

Results Over 1990-2006, no spatial heterogeneity of AL or its subtypes was evidenced. Additionally, none of the most likely clusters identified with SaTScan and FleXScan over the whole period was significant and the systematic search for space-time cluster yielded non-significant results. However, when three sub-periods were considered, five statistically significant non-overlapping spatial clusters were identified.

Conclusion This study did not evidence any global spatial heterogeneity of AL incidence rates in France over the period, 1990-2006. Although no spatial cluster was detected significantly over the whole period, the study identified a few significant spatial clusters in specific periods. Even though the significance levels of those clusters do not strongly support the existence of local risk factors, the clusters may still reflect a slight impact of shared risk factors, including background environmental exposures, which require further investigation.

Key Words Childhood leukaemia; cluster detection; France; spatial clustering/heterogeneity

Author's copy

Introduction

With an incidence of 41.7 cases per million per year in the population of mainland France aged less than 15 years (approximately 11 million children), acute leukaemia (AL) accounts for a third of childhood cancer cases [1]. The French National Registry of Childhood Haematopoietic malignancies (NRCH) has reported that the incidence of acute lymphoblastic leukaemia (ALL), which accounts for 80% of AL cases, is 33.8 cases per million per year and that of acute myeloblastic leukaemia (AML) is 7.1 cases per million per year.

The risk factors for childhood AL have yet to be fully elucidated. In addition to a few genetic factors and exposure to high-dose ionizing radiation and chemotherapy, a few environmental exposures, such as low-dose ionizing radiation, pesticides, road traffic and high-voltage power lines are suspected risk factors. Infection by an unknown common leukaemogenic agent and delayed exposure to common infections have also been suggested as risk factors.

Analysing the spatial distribution of AL incidence is essential for disease surveillance with a view to early detection of emergent localized excesses. Spatial distribution analysis may also contribute to the identification of AL risk factors, including environmental factors. Such studies have been conducted in various countries by several authors (reviewed by McNally and Eden in 2004 [2]). Further studies have been conducted in Great Britain [3] [4] and, more recently, New Zealand [5], Palestine [6], the USA [7], Germany [8] and Denmark [9]. In France, a moderate but borderline statistically significant heterogeneity in the incidence of childhood AL over the period 1990-1994 was observed on the 'Commune' – the smallest administrative units – scale (overdispersion of 1.2%, p=0.06) [10]. In addition, a moderate and borderline statistically significant heterogeneity in AL incidence was observed on the larger 'Département' scale over the period 1990-2004 (overdispersion of 18.6%, p=0.10) [11]. Other results, such as the detection of a large cluster of AML for the period 1995-1999 in an area overlapping a region of high radon concentration [unpublished result, cited in Radiation Protection No. 163, 2009] led to further analyses of AL distribution in France being conducted.

In the present study, the spatial distribution of the incidence of childhood AL in France was studied over a longer period, 1990-2006, on the recent geographic scale of 'living zones' (LZ), the size of which is intermediate between that of the 'Communes' and 'Départements'. From an epidemiological point of view, the LZ scale is pertinent to study of the association between a disease risk and the exposure characteristics of an area in which most of the population spends almost all of its time.

Author's copy

The global spatial heterogeneity of AL incidence (e.g. spatial clustering) in France was evaluated and an investigation of spatial and space-time localized excesses of AL (e.g. clusters) was conducted.

Methods

Cases

Since 1990, the NRCH has registered all the cases of acute leukaemia and lymphoma diagnosed in children aged less than 15 years and living in metropolitan France at the time of diagnosis. Between 1990 and 2006, 7,675 cases of AL were registered in the NRCH: 6,269 cases of ALL and 1,276 cases of AML.

Population

In 1999, the population of mainland France consisted in 58,518,395 people living in 36,565 'Communes', with a median of 380 and a 90th percentile of 2,484 people. The 'Communes' are grouped into 96 'Départements', with median and 90th percentile populations of 503,220 and 1,233,194 people, respectively.

In 2003, a new non-administrative division, the 'living zone' (LZ), was created by the National Institute of Statistics and Economic Studies (INSEE). A LZ, which is comprised of several neighbouring 'Communes', is defined as the smallest territory in which people have access to employment and everyday facilities. There are 1,916 LZ in mainland France (Figure 1). Because some cluster detection methods do not address the particular case of isolated units without any neighbouring units, the 21 LZ located on islands were excluded from the analysis, which thus addressed 1,895 LZ. On average, a LZ consists of 19 'Communes', and has a population of about 30,500 people, of whom 5,450 children (Table 1). The smallest LZ consists of a single 'Commune' with a population of 270 people while the largest LZ has a population of over 9,8 million people (LZ centred on the city of Paris).

The expected number of AL cases was calculated for each 'Commune', age group (< 1 year, 1-4 years, 5-9 years and 10-14 years) and year for the period 1990-2006 under the null hypothesis that childhood AL incidence rates were homogenous throughout the country. The age-specific incidence rates for AL, ALL and AML were provided by the NRCH (Table 2). The person-years at risk for each age group were derived from the populations estimated by INSEE. Over the period 1990-2006, less than one case of AL was expected in 707 of the 1,895 LZ, with an overall mean of 4.1 per LZ, and a maximum of 1,400 cases for the Paris LZ (Table 1).

Spatial heterogeneity methods

Author's copy

The global spatial heterogeneity in the data was investigated using the Potthoff and Whittinghill (P-W) test for overdispersion [12] and the Rogerson statistic [13] derived from Tango's test for heterogeneity. The P-W method assumes that the number of cases in each LZ is Poisson distributed under the null hypothesis of homogeneous incidence rates (H0) and follows a negative binomial distribution under the alternative hypothesis that observed cases do have the mean expected under H0 but are over-dispersed with a variance to the mean ratio equal to $1+\beta$, in which $\beta>0$.

The Rogerson's statistic R [13], which is equal to 0 under the null hypothesis of no spatial heterogeneity, combines a chi-squared statistic of goodness-of-fit and a statistic of spatial autocorrelation between neighbouring units. The neighbourhood limit was set to 40 km. Analyses were performed with the R[1] (v. 2.10.1) software.

Cluster detection methods

Three cluster detection methods sharing the likelihood ratio approach developed by Kulldorff and Nagarwalla [14] were used. The methods scan the territory by creating collections cluster candidates composed of neighbouring LZ and consider, for each, the likelihood ratio based on the alternative hypothesis that the incidence is higher inside than outside the cluster candidate and the null hypothesis that the incidence rates are equal. The most likely cluster (MLC) is then the candidate area that maximizes the likelihood ratio function.

The spatial SaTScan method with circular moving windows (SaTS-c) [14] was used with the maximum size set to 100 LZ, and the method extension with elliptic windows [15], with or without a penalty introduced to reduce the likelihood of the most eccentric cluster candidates (SaTS-e0 and SaTS-e1, respectively). The analyses were performed with the SaTScan (v. 9.0.1)[2] software.

The FleXScan method [16], which aims to detect arbitrarily shaped clusters, was also applied, using the FleXScan (v3.0)[3] software. The method is free from the regular shape constraints and considers all the connected zones included in a given neighbourhood of each LZ as cluster candidates. A restriction can be added to retain only areas made-up of high-risk LZ. Windows of at most 25 LZ were considered in the present study with (FleXS-r(25)) or without (FleXS(25)) a restriction. A 100 LZ window was also considered for FleXScan but only with restriction because of the running time (FleXS-r(100)).

[1] R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
[2] SaTScanTM is a trademark of Martin Kulldorff. The SaTScanTM software was developed under the joint auspices of Martin Kulldorff, the National Cancer Institute, and Farzad Mostashari of the New York City Department of Health and Mental Hygiene.
[3] Takahashi K, Yokoyama T and Tango T. FleXScan: Software for the Flexible Scan Statistic. National Institute of Public Health, Japan, 2005-2006.

Author's copy

The space-time SaTScan method [17] was used for detection of space-time clusters.

The p-value of the tests were evaluated using 999 Monte Carlo simulations under the null hypothesis of homogeneous incidence rates for each application. If the p-value was smaller than 0.1, the MLC was considered significant.

The analyses were conducted over the whole time period and on three sub-periods, 1990-1994, 1995-2000, 2001-2006, for all AL and the ALL and AML.

Results

Spatial heterogeneity

The P-W test did not show any statistically significant overdispersion of AL, ALL or AML cases. The estimated overdispersion parameter β was equal to -0.04 (p=0.91, Table 3). The Rogerson's test did not evidence any significant heterogeneity in AL incidence rates (R = 0.25, p=0.39, Table 3) or in ALL or AML incidence rates. The results were unchanged when the neighbourhood limit was set to 20, 30, 50 and 60 km instead of 40 km (data not shown).

Cluster detection

Over the period 1990-2006, none of the MLC identified with the spatial SaTScan and FleXScan methods was significant overall or for ALL or AML (Table 4). No significant cluster was identified using the space-time SaTScan method (Table 5).

For the 3 sub-periods, four of the MLC detected by the spatial SaTScan method were statistically significant (Table 6, Figure 2). For the period 1990-1994, a cluster of AL was detected by SaTS-e0 and SaTS-e1 (cluster A), but was only significant with SaTS-e1 (p=0.08). The cluster was also included in a non-significant MLC identified with the circular window (SaTS-c) on that period. The cluster corresponded exactly to the geographic basis of the non-significant MLC found with the space-time SaTScan method for the period 1990-1995. For 1995-2000, a large AML cluster was detected with SaTS-e0 (cluster B, p=0.08), which partially overlapped the non-significant MLC detected with SaTS-e1 and SaTS-c. An ALL cluster was detected with SaTS-c and SaTS-e1 (cluster C, p=0.06 and p=0.08, respectively) with 24 out of 64 LZ included in the non-significant MLC detected by SaTS-e0. The cluster was also detected in the space-time analysis as the ALL non-significant MLC for the shorter period 1995-1999. A cluster of ALL cases aged 1-4 years old was evidenced by SaTS-e1 (cluster

Author's copy

D, p=0.09). It was also the non-significant MLC with SaTS-e0 and comprised the non-significant MLC identified with SaTS-c. The FleXScan method with a restriction detected a cluster of 1-4 year ALL cases over 2001-2006, cluster E (p=0.08), with the maximum window size set to 25 LZ. An overlapping 13-LZ cluster (p=0.05) was also detected with the maximum window size set to 100 LZ (not shown). Cluster E was included in the non-significant MLC identified with the spatial SaTScan method for 2001-2006 and with the space-time SaTScan method for 2004-2006. As shown in Figure 2, the five significant clusters did not overlap. Although the FleXScan method did not detect any other statistically significant cluster, the method identified non-significant MLC in areas in which the SaTScan method had detected significant clusters (A, B, C and D).

None of the significant clusters identified in a given period remained statistically significant when the Bonferroni correction was applied to account for the 3 time periods considered and 18 tests carried out for each AL type.

Discussion

This study of the spatial variability of the French AL, ALL and AML incidence rates over the period 1990-2006 on the LZ scale did not evidence any global spatial heterogeneity using the Potthoff-Whittinghill [12] or Rogerson methods [13]. In the investigation of potential spatial clusters for the whole period, 1990-2006, neither the SaTScan method nor the FleXScan method detected significant clusters of acute leukaemia cases. The systematic search for space-time clusters over 1990-2006 resulted in non-significant clusters. However, when 3 periods (1990-1994, 1995-2000 and 2001-2006) were considered, five statistically significant spatial clusters were detected.

Epidemiological investigation for new possible environmental risk factors may adopt various complementary approaches. In order to collect information on the association between a given exposure and a given disease, it is appropriate to design studies focused on the exposure. Depending on the amount of prior knowledge available, studies, such as those conducted by the NRCH team, may preferably be ecological [18], [19] or individual, case-control [20] or cohort studies. It may also be of interest to examine whether variability of incidence rates exists on a given geographic scale and, if so, to try to identify a possible exposure that correlates with the spatial variations. Analyzing the global spatial heterogeneity of the incidence of a disease and localizing potential local excesses may therefore be informative.

Author's copy

In the 17-year period, 7,675 cases of childhood acute leukaemia (AL) were registered in France by the NRCH, which has been officially recognized for the high level and quality of its case ascertainment (estimated exhaustiveness: 99.4% [21]; about three sources per case on average [1]). The analyses were conducted for all AL, ALL and AML. Each of the subtypes encompasses other subtypes, which may be associated with different risk factors. With regard to ALL, the study focused on cases aged 1-4 years, 78% of which were precursor B-cell ALL. The numbers of cases in the AML subtypes were too small for separate analyses. Thus, any spatial heterogeneity or localized cluster possibly due to an underlying risk factor specific to a given AML or ALL subtype, other than precursor B-cell ALL, would not have been evidenced by the present study.

Scale and boundaries are determinants in spatial heterogeneity analyses since the geographic scale on which the spatial heterogeneity is investigated may obscure a spatial heterogeneity that may exist on a larger scale, such as on the 'Département' scale in France [11], or hinder detection of a smaller scale heterogeneity, such as on the 'Commune' scale in France [10]. In addition, clusters may cross census boundaries and thus be less likely to be identified. In this study, the analyses were carried out on the 'living zone' (LZ) scale. LZ are zones with homogeneous activities and are structured around the daily migratory flux of the population. The population of these non-administrative units also has a higher level of geographic, social, cultural and economic homogeneity. This scale is appropriate for the detection of excesses spread over large areas and compatible with exposures shared by most of the LZ population, including environmental exposures. However, this scale is not pertinent for the detection of clusters as small as the frequently-reported suspected clusters in a school or block of houses.

This study did not identify any spatial or spatiotemporal local clusters over the whole period 1990-2006. Because of the great heterogeneity of LZ population size, the maximum cluster size was specified as a number of LZ instead of a given proportion of the total population, e.g. 50%, as is often done by other authors. Since this study was conducted nationwide, a maximum window size of 10% of the population would have been a reasonable choice. However, with that limit, four LZ, including the Paris LZ, and all the potential windows centred on them would not have been considered cluster candidates. In contrast, a higher limit might have generated excessively large clusters.

The significant spatial clusters identified in the analyses by period were specific to different subtypes of AL, located in disjointed areas and observed in different periods (AL 1990-1994; AML 1995-2000 similar to the

Author's copy

cluster previously found over the period 1995-1999 [unpublished result, cited in Radiation Protection No. 163, 2009]; ALL 1995-2000; ALL 1-4 years 1995-2000 and 2001-2006).

In all the situations, the most likely clusters identified by the SaTScan and FleXScan methods were located in the same region and, except for AML, the significant clusters also coincided with the geographic basis of the non-significant most likely cluster in the systematic space-time analysis. The results of the SaTScan and FleXScan methods were thus similar with regard to the identification of the most likely cluster areas, i.e. the areas associated with the maximum likelihood ratios, but the methods were not concordant with regard to statistical significance.

In the cluster detection methods, a cluster candidate is identified as the most likely cluster, among all the candidate zones, if it maximizes the likelihood ratio function. The p-value increases with the number of candidate zones considered, which is closely related to the size of the study area and the size of the scanning window [22]. Furthermore, the power of cluster detection methods depends on the spatial distribution of the geographic units and their population size, as has been previously suggested ([23],[24],[25]). Recently, a power study based on the data from the NRCH was conducted in order to evaluate the performance of the SaTScan and FleXScan methods and three other recent methods ([26] and M. Costa, submitted for publication) under several realistic alternative scenarios [27]. In the situations where the 'true' cluster was of medium or large size (i.e. more than 45 cases of AL expected in 1990-2006) and the risk ratio greater than or equal to 2, all the methods had a usual power greater than 90% and the SaTS-e0 and FleXS windows had the highest sensitivity. If such a cluster existed in France, it is thus likely that it would be significantly detected by both the SaTScan and FleXScan methods. The power study also showed that clusters with 20 expected cases and a risk ratio equal to 1.5 or 2 were almost undetectable by any of the methods. The low power may explain the discordance between the significant results of the SaTScan and the FleXScan methods in the present study, since all but one of the significant clusters found consisted in less than 20 expected cases.

To the authors' knowledge, few significant clusters of rare diseases, such as childhood acute leukaemia, have been detected with cluster detection methods. Among the 8 studies that used the SaTScan method ([6], [7], [8], [10], [14], [22], [28], [29]), only three detected significant spatial or space-time clusters ([7], [14], [22]). No cause was identified.

Author's copy

It is noteworthy that all the significant clusters detected for the period before 2001 (clusters A, B, C and D) were located in the East of France where the radioactive fallout (Caesium 137) from the Chernobyl accident was greatest [30]. To a lesser extent, the clusters also seem to be located in areas with higher radon exposure. These findings call for further investigation.

In conclusion, this study, which addressed the whole of mainland France over the period 1990-2006, did not evidence any spatial heterogeneity of childhood AL incidence rates. However, the study identified a few significant spatial clusters within specific periods. Although the significance levels of the clusters do not strongly support the existence of localized risk factors, the clusters may still reflect a slight impact of risk factors shared on the LZ scale, including background environmental exposures, which require further investigation.

Author's copy

Acknowledgements

This work was supported by grants from INSERM, InVS, Fondation de France and EDF. The authors are grateful to A. Mullarky for his skilful revision of the manuscript.

Author's copy

References

1. Lacour B, Guyot-Goubin A, Guissou S, Bellec S, Desandes E, Clavel J. Incidence of childhood cancer in France: National Children Cancer Registries, 2000-2004. Eur J Cancer Prev. 2010;19(3):173-81.

2. McNally RJ, Eden TO. An infectious aetiology for childhood acute leukaemia: a review of the evidence. Br J Haematol. 2004;127(3):243-63. doi:10.1111/j.1365-2141.2004.05166.x.

3. McNally RJ, Alexander FE, Bithell JF. Space-time clustering of childhood cancer in great Britain: a national study, 1969-1993. Int J Cancer. 2006;118(11):2840-6. doi:10.1002/ijc.21726.

4. McNally RJ, Alexander FE, Vincent TJ, Murphy MF. Spatial clustering of childhood cancer in Great Britain during the period 1969-1993. Int J Cancer. 2009;124(4):932-6. doi:10.1002/ijc.23965.

5. Dockerty JD, Sharples KJ, Borman B. An assessment of spatial clustering of leukaemias and lymphomas among young people in New Zealand. J Epidemiol Community Health. 1999;53(3):154-8.

6. Bailony MR, Hararah MK, Salhab AR, Ghannam I, Abdeen Z, Ghannam J. Cancer registration and healthcare access in West Bank, Palestine: A GIS analysis of childhood cancer, 1998-2007. Int J Cancer. 2010. doi:10.1002/ijc.25732.

7. Amin R, Bohnert A, Holmes L, Rajasekaran A, Assanasen C. Epidemiologic mapping of Florida childhood cancer clusters. Pediatr Blood Cancer. 2010;54(4):511-8. doi:10.1002/pbc.22403.

8. Schmiedel S, Blettner M, Kaatsch P, Schuz J. Spatial clustering and space-time clusters of leukemia among children in Germany, 1987-2007. Eur J Epidemiol. 2010;25(9):627-33. doi:10.1007/s10654-010-9488-7.

9. Schmiedel S, Jacquez GM, Blettner M, Schuz J. Spatial clustering of leukemia and type 1 diabetes in children in Denmark. Cancer Causes Control. 2011. doi:10.1007/s10552-011-9755-2.

10. Bellec S, Hemon D, Rudant J, Goubin A, Clavel J. Spatial and space-time clustering of childhood acute leukaemia in France from 1990 to 2000: a nationwide study. Br J Cancer. 2006;94(5):763-70. doi:10.1038/sj.bjc.6602980.

11. Faure C, Mollie A, Bellec S, Guyot-Goubin A, Clavel J, Hemon D. Geographical variations in the incidence of childhood acute leukaemia in France over the period 1990-2004. Eur J Cancer Prev. 2009;18(4):267-79. doi:10.1097/CEJ.0b013e32832bf43a.

12. Potthoff RF, Whittinghill M. Testing for homogeneity. II. The Poisson distribution. Biometrika. 1966;53(1):183-90.

13. Rogerson PA. The Detection of Clusters Using a Spatial Version of the Chi-Square Goodness-of-Fit Statistic.

14. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. Stat Med. 1995;14(8):799-810.

15. Kulldorff M, Huang L, Pickle L, Duczmal L. An elliptic spatial scan statistic. Stat Med. 2006;25(22):3929-43.

16. Tango T, Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. Int J Health Geogr. 2005;4:11.

17. Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. Am J Public Health. 1998;88(9):1377-80.

18. Evrard AS, Hemon D, Billon S, Laurier D, Jougla E, Tirmarche M et al. Ecological association between indoor radon concentration and childhood leukaemia incidence in France, 1990-1998. Eur J Cancer Prev. 2005;14(2):147-57.

19. Evrard AS, Hemon D, Billon S, Laurier D, Jougla E, Tirmarche M et al. Childhood leukemia incidence and exposure to indoor radon, terrestrial and cosmic gamma radiation. Health Phys. 2006;90(6):569-79. doi:10.1097/01.hp.0000198787.93305.35.

20. Amigou A, Sermage-Faure C, Orsi L, Leverger G, Baruchel A, Bertrand Y et al. Road traffic and childhood leukemia: the ESCALE study (SFCE). Environ Health Perspect. 2011;119(4):566-72. doi:10.1289/ehp.1002429.

21. Clavel J, Goubin A, Auclerc MF, Auvrignon A, Waterkeyn C, Patte C et al. Incidence of childhood leukaemia and non-Hodgkin's lymphoma in France: National Registry of Childhood Leukaemia and Lymphoma, 1990-1999. Eur J Cancer Prev. 2004;13(2):97-103.

Author's copy

22. Wheeler DC. A comparison of spatial clustering and cluster detection techniques for childhood leukemia incidence in Ohio, 1996-2003. Int J Health Geogr. 2007;6:13. doi:10.1186/1476-072x-6-13.

23. Song C, Kulldorff M. Power evaluation of disease clustering tests. Int J Health Geogr. 2003;2(1):9. doi:10.1186/1476-072x-2-9.

24. Huang L, Pickle LW, Das B. Evaluating spatial methods for investigating global clustering and cluster detection of cancer cases. Stat Med. 2008;27(25):5111-42. doi:10.1002/sim.3342.

25. Waller LA, Hill EG, Rudd RA. The geography of power: statistical performance of tests of clusters and clustering in heterogeneous populations. Stat Med. 2006;25(5):853-65. doi:10.1002/sim.2418.

26. Duczmal L, Cançado ALF, Takahashi RHC, Bessegato LF. A genetic algorithm for irregularly shaped spatial scan statistics. Computational Statistics & data Analysis. 2007;52:43-52.

27. Goujon-Bellec S, Demoury C, Guyot-Goubin A, Hemon D, Clavel J. Detection of clusters of a rare disease over a large territory: performance of cluster detection methods. International Journal of Health Geographics. 2011;10(1):53.

28. Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N. Childhood leukaemia in Sweden: using GIS and a spatial scan statistic for cluster detection. Stat Med. 1996;15(7-9):707-15.

29. Mosavi-Jarrahi A, Moini M, Mohagheghi MA, Alebouyeh M, Yazdizadeh B, Shahabian A et al. Clustering of childhood cancer in the inner city of Tehran metropolitan area: a GIS-based analysis. Int J Hyg Environ Health. 2007;210(2):113-9. doi:10.1016/j.ijheh.2006.08.005.

30. Verger P C-CL. Evaluation des conséquences sanitaires de l'accident de Tchernobyl en France: dispositif de surveillance épidémiologique, état des connaissances, évaluation des risques et perspectives: IPSN, InVS 2009.

Author's copy

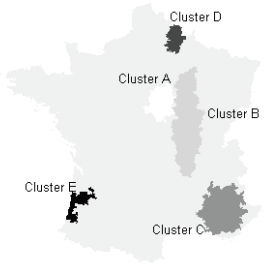
Fig. 1 The 1,916 'living zones' (LZ) in France



LZ is a geographic unit created in 2003 by the National Institute of Statistics and Economic Studies (INSEE); it is the smallest territory in which residents have access to both everyday facilities and employment. *(Source: Géocible)*

Author's copy

Fig. 2 Significant spatial clusters of acute leukemia (AL) detected in France between 1990 and 2006, with two cluster detection methods [14-16]



Cluster A: detected with the SaTS-e1 method, AL for the period 1990-1994, p = 0.08

Cluster B: detected with the SaTS-e0 method, AML for the period 1995-2000, p = 0.08

Cluster C: detected with the SaTS-c and SaTS-e1 method, ALL for the period 1995-2000, p = 0.06 and p = 0.08

Cluster D: detected with the SaTS-e1 method, ALL 1-4 years for the period 1995-2000, p = 0.09

Cluster E: detected with the FlexS-r(25) method, ALL 1-4 years for the period 2001-2006, p = 0.08

See §results and Table 6 for more details

Author's copy

Table 1 Area and composition of the French Living Zones (LZ), in terms of number of 'Communes', population size and expected number of children acute leukaemia (AL) cases

	Mean	Minimum	Q1 [a]	Median	Q3 [b]	P90 [c]	Maximum
Area (km^2)	282.3	0.4	108.8	193.0	340.2	629.4	3,863.2
Number of 'Communes'	19.1	1	7	13	24	41	556
Estimated total population (1999)	30,542.0	270.0	6,219.8	9,754.5	17,968.0	44,514.5	9,802,327.0
Estimated population [0-14 years] (1999)	5,453.7	71.0	1,068.0	1,725.0	3,298.0	7,951.5	1,823,195.0
Expected number of LA cases	4.1	0.1	0.8	1.3	2.4	5.8	1,388.3

[a] first quartile, [b] third quartile, [c] 90th percentile

Author's copy

Table 2 Incidence rates of AL and its sub-types in France for the periods 1990-1994, 1995-2000 and 2001-2006

Incidence rate (per 100 000 person-year)		1990-1994		1995-2000		2001-2006	
		Crude	Age standardized	Crude	Age standardized	Crude	Age standardized
All AL[a]		3,93	4,13	4,00	4,23	4,16	4,34
AML[b]		0,66	0,68	0,70	0,72	0,66	0,67
ALL[c]	All ages	3,22	3,39	3,25	3,45	3,41	3,57
	1-4 years	5,91	-	5,96	-	6,05	-

[a] AL: acute leukaemia; [b] AML: acute myeloblastic leukaemia; [c] ALL: acute lymphoblastic leukaemia

Author's copy

Table 3 Global spatial heterogeneity of childhood acute leukaemia incidence rates in France over the period 1990-2006

	P-W method [a]		Rogerson's method [b]	
	$\hat{\beta}$	p-value	Ṙ	p-value
All AL [c] (n = 7675)	-0.04	0.91	0.25	0.39
AML [d] (n = 1276)	-0.04	0.90	1.66	0.21
ALL [e] (n = 6269)	-0.03	0.86	0.32	0.30
ALL [e] 1-4 years (n = 3013)	-0.05	0.93	0.57	0.74

[a] Potthoff-Whittinghill method [13]. $\hat{\beta}$: estimated value of the overdispersion parameter ß under the hypothesis that the number of cases follows a negative binomial distribution and do have the mean expected under H0 but are over-dispersed with a variance to the mean ratio equal to 1+β, in which β>0 (H0: β = 0);

[b] Rogerson's method with a 40-km neighbourhood Ṙ [14]. Ṙ: estimated value of the test statistic R(H0: R = 0);

[c] AL: children acute leukaemia, n = number of cases; [d] AML: children acute myeloblastic leukaemia; [e]ALL: children acute lymphoblastic leukaemia; n = number of cases.

Author's copy

Table 4 Detection of spatial cluster of childhood acute leukaemia in France over the period 1990-2006

Method			SaTScan			FleXScan		
	Window shape		Circular	Elliptic		Flexible		
				without penalty	with penalty	without restriction	with restriction	with restriction
	Maximum cluster size		100	100	100	25	25	100
All AL [a]		O [d]	263	222	256	46	33	104
		E [e]	212.1	167.2	200.0	23.1	15.5	61.4
		LZ [f]	73	42	67	10	5	14
		p [g]	0.93	0.74	0.68	0.89	0.80	0.20
AML [b]		O [d]	14	47	60	15	13	13
		E [e]	4.8	23.8	34.9	3.4	2.9	2.9
		LZ [f]	11	71	85	9	6	6
		p [g]	0.86	0.50	0.55	0.36	0.31	0.51
ALL [c]	All ages	O [d]	244	250	219	34	26	57
		E [e]	187.9	187.0	163.2	13.3	9.7	30.1
		LZ [f]	70	71	67	10	6	10
		p [g]	0.23	0.27	0.23	0.35	0.36	0.54
	1-4 years	O [d]	9	113	9	16	15	15
		E [e]	2.1	76.0	2.1	4.3	4.4	4.4
		LZ [f]	6	50	6	12	5	5
		p [g]	0.74	0.79	0.88	0.76	0.74	0.86

[a] AL: acute leukaemia; [b] AML: acute myeloblastic leukaemia; [c] ALL: acute lymphoblastic leukaemia

[d] O = number of case observed within the most likely cluster identified (MLC); [e] E = number of cases expected within the MLC; [f] LZ= number of 'living zones' within the MLC

[g] p = statistical level of significance of the MLC

Author's copy

Table 5 Detection of space-time clusters of childhood acute leukaemia in France over the period 1990-2006 with the space-time SaTScan method

Window shape (with a maximum cluster size of 100 LZ)			Circular	Elliptic	
				without penalty	with penalty
All AL [a]		O [d]	52	51	51
		E [e]	25.2	23.0	23.0
		LZ [f]	34	28	28
		Period	1990-1995	1990-1995	1990-1995
		p [g]	0.57	0.77	0.33
AML [b]		O [d]	4	8	4
		E [e]	0.1	0.7	0.1
		LZ [f]	4	4	3
		Period	2004-2005	1991-1995	2004-2005
		p [g]	0.71	0.78	0.81
ALL [c]	All ages	O [d]	122	111	122
		E [e]	76.3	66.4	76.3
		LZ [f]	64	67	64
		Period	1995-1999	1997-2003	1995-1999
		p [g]	0.36	0.79	0.56
	1-4 years	O [d]	11	14	14
		E [e]	2.0	2.7	2.7
		LZ [f]	13	20	20
		Period	2004-2006	2004-2006	2004-2006
		p [g]	0.82	0.96	0.77

[a] AL: acute leukaemia; [b] AML: acute myeloblastic leukaemia; [c] ALL: acute lymphoblastic leukaemia

[d] O = number of case observed within the most likely cluster identified (MLC); [e] E = number of cases expected within the MLC; [f] LZ= number of 'living zones' within the MLC

[g] p = statistical level of significance of the MLC

Author's copy

Table 6 Detection of spatial clusters of acute leukaemia in France for the sub-periods 1990-1994, 1995-2000 and 2001-2006

Method		Most likely cluster identified for the period 1990-1994					Most likely cluster identified for the period 1995-2000					Most likely cluster identified for the period 2001-2006				
		SaTScan			FleXScan		SaTScan			FleXScan		SaTScan			FleXScan	
Window shape		Circular	Elliptic penalty		Flexible restriction		Circular	Elliptic penalty		Flexible restriction		Circular	Elliptic penalty		Flexible restriction	
			No	Yes	No	Yes		No	Yes	No	Yes		No	Yes	No	Yes
Maximum cluster size		100	100	100	25	25	100	100	100	25	25	100	100	100	25	25
All AL [a]	O [d]	43	42	42	26	24	7	14	7	34	8	123	16	54	26	14
	E [e]	21.0	19.1	19.1	8.8	7.9	1.1	3.4	0.8	12.8	1.1	87.0	4.6	30.4	9.4	3.3
	LZ [f]	34	28	28	10	7	4	9	3	17	4	79	7	31	11	9
	p [g] (cluster name)	0.10	0.21	**0.08(A)**	0.35	0.17	0.44	0.48	0.24	0.20	0.43	0.54	0.68	0.53	0.60	0.32
AML [b]	O [d]	8	16	6	8	10	53	41	51	10	12	5	10	5	7	7
	E [e]	1.5	3.9	0.7	1.0	1.9	29.5	18.1	26.6	1.8	3.3	0.4	1.8	0.3	0.8	0.8
	LZ [f]	7	21	9	8	8	97	96	96	9	4	4	22	3	5	5
	p [g] (cluster name)	0.46	0.12	0.49	0.41	0.20	0.16	**0.08(B)**	0.11	0.61	0.64	0.20	0.35	0.22	0.47	0.12
ALL [c] All ages	O [d]	7	81	7	26	23	136	139	136	29	87	111	108	116	21	13
	E [e]	1.1	49.4	1.1	10.1	8.6	91.7	91.4	91.7	10.8	54.0	77.8	69.6	77.5	7.0	3.4
	LZ [f]	2	69	2	11	8	64	48	64	16	11	97	70	76	10	5
	p [g] (cluster name)	0.44	0.56	0.54	0.87	0.55	**0.06(C)**	0.11	**0.08(C)**	0.43	0.45	0.78	0.44	0.49	0.82	0.68
ALL [c] 1-4 years	O [d]	15	7	5	14	14	18	25	25	19	12	62	67	62	17	17
	E [e]	5.0	0.9	0.5	3.5	3.5	5.9	8.7	8.7	4.7	2.6	37.0	38.1	37.0	4.1	4.1
	LZ [f]	18	7	5	7	7	20	24	24	13	10	75	76	75	10	10
	p [g] (cluster name)	0.57	0.61	0.63	0.71	0.28	0.22	0.23	**0.09(D)**	0.11	0.32	0.37	0.33	0.55	0.23	**0.08(E)**

[a] AL: acute leukaemia; [b] AML: acute myeloblastic leukaemia; [c] ALL: acute lymphoblastic leukaemia; [d] O = number of case observed within the most likely cluster identified (MLC); [e] E = number of cases expected within the MLC; [f] LZ= number of 'living zones' within the MLC; [g] p = statistical level of significance of the MLC

Authors copy

A statistical evaluation of the influence of housing characteristics and geogenic radon potential on indoor radon concentrations in France

DEMOURY C.[*a,b], IELSCH G.[*c], HEMON D.[a,b], LAURENT O.[d], LAURIER D.[d], CLAVEL J.[§a,b], GUILLEVIC J [§c].

[*] Joint first authors. The authors consider that the first two authors should be regarded as joint First Authors.

[§] Joint last authors. The authors consider that the last two authors should be regarded as joint Last Authors.

[a] Inserm, CESP Center for Research in Epidemiology and Population Health, U1018, Environmental epidemiology of cancer Team, 16 avenue Paul Vaillant-Couturier, F-94807, Villejuif, France

[b] Univ Paris-Sud, UMRS 1018, 16 avenue Paul Vaillant-Couturier, F-94807, Villejuif, France

[c] Institut de Radioprotection et de Sûreté Nucléaire, Bureau d'étude et d'expertise sur la radioactivité naturelle. IRSN, PRP-DGE, SEDRAN, BRN, BP17, 92262 Fontenay aux Roses Cedex, France

[d] Institut de Radioprotection et de Sûreté Nucléaire, Laboratoire d'épidémiologie. IRSN, PRP-HOM, SRBE, LEPID, BP17, 92262 Fontenay aux Roses Cedex, France

Email addresses:

claire.demoury@insem.fr (C. Demoury), geraldine.ielsch@irsn.fr (G. Ielsch), denis.hemon@inserm.fr (D. Hémon), olaurent@uci.edu (O. Laurent), dominique.laurier@irsn.fr (D. Laurier), jacqueline.clavel@inserm.fr (J. Clavel), jerome.guillevic@irsn.fr (J. Guillevic)

Corresponding authors:

For the final accepted version of the paper: Géraldine IELSCH

Email: geraldine.ielsch@irsn.fr;Phone: +33 1 58 35 80 81; Fax: +33 1 58 35 80 35

Funding

This study was supported in part by grant EST-09-10 from the French Agency for Food, Environmental and Occupational Health & Safety (ANSES) and by the Institute for Radiological Protection and Nuclear Safety. ANSES had no role in study design, the collection, analysis and interpretation of the data, writing the article or the decision to submit it for publication.

Author's copy

Abstract

Radon-222 is a radioactive natural gas produced by the decay of radium-226, known to be the main contributor to natural background radiation exposure. Effective risk management needs to determine the areas in which the density of buildings with high radon levels is likely to be highest. Predicting radon exposure from the location and characteristics of a dwelling could also contribute to epidemiological studies. Beginning in the nineteen-eighties, a national radon survey consisting in more than 10,000 measurements of indoor radon concentrations was conducted in French dwellings by the Institute for Radiological Protection and Nuclear Safety (IRSN). Housing characteristics, which may influence radon accumulation in dwellings, were also collected. More recently, the IRSN generated a French geogenic radon potential map based on the interpretation of geological features. The present study analyzed the two datasets to investigate the factors influencing indoor radon concentrations using statistical modeling and to determine the optimum use of the information on geogenic radon potential that showed the best statistical association with indoor radon concentration. The results showed that the variables associated with indoor radon concentrations were geogenic radon potential, building material, year of construction, foundation type, building type and floor level. The model, which included the surrounding geogenic radon potential (i.e. the average geogenic radon potential within a disc of radius 20 km centered on the indoor radon measurement point) and variables describing house-specific factors and lifestyle explained about 20% of the overall variability of the logarithm of radon concentration. The surrounding geogenic radon potential was fairly closely associated with the local average indoor radon concentration. The prevalence of exposure to radon above specific thresholds and the average exposures to radon clearly increased with increasing classes of geogenic radon potential. Combining the two datasets enabled improved assessment of radon exposure in a given area in France.

Keywords: Indoor radon; Geology; Radon mapping; Statistical modeling; France.

Abbreviations:

RC: indoor radon concentration

GRP: geogenic radon potential

IRSN: Institute for Radiological Protection and Nuclear Safety

logRC: decimal logarithms of the RC

Author

1. Introduction

The exposure of the French population to ionizing radiation is mainly due to natural radiation, particularly inhalation of radon decay products. The population exposure to radon is highly variable over the country with indoor radon concentrations varying from a few Bq/m^3 to several thousand Bq/m^3. In order to assess the variability, various methods may be developed and tested. The methods may combine several indicators, such as indoor radon concentration or soil gas radon concentration measurements, soil properties and geological data. In France two national datasets were built by the Institute for Radiological Protection and Nuclear Safety (IRSN) to identify radon prone areas and assist the authorities in radon risk management. The first dataset contains the results of indoor radon measurements in dwellings and the second is a map of geogenic radon potential.

Beginning in the nineteen-eighties, a French nationwide indoor radon survey was carried out and included over 10,000 indoor radon concentration (RC) (Gambard et al., 2000) measurements. Housing characteristics and information on the lifestyle of the dwellings' residents were also collected during the survey. Two studies of this dataset were conducted. The first (Verger et al., 1994) consisted in a statistical analysis of the building characteristics potentially affecting RC and was based on 1,548 indoor radon results for 15 *départements* (French administrative areas). The second study (Billon et al., 2005) enabled the estimates of the French population's exposure to natural radiation to be updated.

Several international studies have also shown the contribution of geology to indoor radon concentrations (Apte et al., 1999; Ielsch et al., 2002; Ielsch et al., 2001; Miles and Appleton, 2005; Sundal et al., 2004; Zhu et al., 2001; Zhu et al., 1998) as the dominant source of indoor radon is the ground on which a dwelling is built. In France, the IRSN has conducted several studies on the influence of geological factors on radon exhalation rates and indoor radon concentrations (Ielsch et al., 2001; Ielsch et al., 2002; Ielsch and Cuney, 2004). The map of geogenic radon potential (GRP) in France was compiled in 2010 and consisted of determining the ability of geological formations to produce radon and facilitate its transfer into the atmosphere (Ielsch et al., 2010).

Other recent studies, carried out in various countries, have used statistical models to assess indoor radon exposure. The models combined different parameters such as housing characteristics, soil properties, geology and gamma dose rate (Andersen et al., 2007; Appleton and Miles, 2010;

Author's copy

Friedmann and Groller, 2010; Hauri et al., 2012; Hunter et al., 2009; Kemski et al., 2009; Makelainen et al., 2001).

The purpose of the present study was to describe the statistical associations between GRP, house-specific factors, lifestyle factors and RC on the basis of the two national datasets available in France. The ability of the various variables to assess radon exposure in France has been evaluated and discussed. The results may enable improvement of radon mapping in France, and thus be of assistance to the authorities for radon risk management and provide useful exposure data for epidemiological studies.

2. Materials and methods

2.1 Database of indoor radon concentration measurements

In France, the national indoor radon measurement survey was carried out from 1982 to 2003 by the IRSN in cooperation with the Ministry of Health. The main objectives were to identify radon-prone areas in France, estimate the percentage ofdwellings above high radon levels (for example 100 and 400 Bq/m^3), and investigate for the factors affecting radon concentrations (Billon et al., 2005; Gambard et al., 2000). RC was measured in dwellings (mainly houses) using LR 115 track-etch detectors (Andru, 1998), placed in the main room (living room or kitchen, generally on the ground floor) and exposed for at least 2 months (and for less than one year), in winter or early spring. The measurements made on underground floors (e.g. cellars) were excluded. The measurement uncertainty varied from 37% for concentrations below 10 Bq/m^3 to <8% for concentrations above 1,000 Bq/m^3. The radon detectors were placed and collected by volunteers recruited through contacts in local health authorities. The volunteers also completed a questionnaire for each measurement location to record factors liable to influence RC, including year of building construction, building materials, type of foundations, building type, room type, floor on which the measurement was made, air renewal frequency and type of ventilation.

Out of the 10,843 RC results in the database, 7,699 were geocoded at the dwelling address. The NAVTEQ and IGN databases were used to generate coordinates with a precision of between 20 m (for

Author's copy

42% of the measurements) and 250 m. Most of the coordinates (i.e. 90%) were geocoded with a precision of 100 m or better. For the other 3,144 dwellings, the only available address information was the *commune* (i.e. the smallest French administrative unit; mainland France is made up of 36,570 *communes*). In those cases, the spatial coordinates of the town hall of the *commune* were assigned to the dwellings. It is well known that indoor radon concentration shows seasonal variations during the year. Considering that measurements were made during a period of generally 2−3 months, seasonal correction factors, estimated using Baysson's methodology (Baysson et al., 2003), were applied to the RC. The resulting values, which are estimates of the annual means, were used in the study.

2.2 Geogenic radon potential map

In 2009-2010, following a request from the French Nuclear Safety Authority, the IRSN developed a method for generating a GRP map of France (Ielsch et al., 2010) with a view to contributing to radon risk management. The method consisted in classifying the geological units in terms of their ability to produce radon and facilitate its exhalation at the soil surface. It was based on the interpretation of the existing national geological data and their cartographic analysis using a geographic information system. The geological formations were classified by uranium content in order to create a potential radon source map. The initial map was then improved by taking into account the main additional parameters, such as faults, underground mines (galleries and shafts) and thermo-mineral groundwater sources, which control the preferential pathways of radon through the ground and which may increase radon levels in the soil. GRP was expressed using 5 qualitative classes: low, low to moderate (heterogeneous), moderate, high, and very high. A class was assigned to each geological feature defined by the boundaries of the geological map of France (1:1,000,000). The map provides a general trend in the variation of GRP on the scale considered.

2.3 Geogenic radon potential quantification

First, the RC measurement sites were superimposed on the GRP map in order to assign a 'pinpoint' GRP class 'k' (k = 1;…; 5) to each of the 10,843 dwellings in which RC was measured. A quantitative score 'S_k' was then allocated to each of the five ordinal GRP classes. 'S_k' was calculated as the

Author's copy

arithmetic mean of the decimal logarithms of the RC (logRC) that were measured in dwellings located in areas where the GRP was classified in category 'k'. A 'pinpoint' GRP score was thus assigned to each site where RC was measured.

Additionally, the average values of the GRP scores ('averaged' GRP) were determined in circular buffers centered on the RC measurement points (Fig. 1.a). Considering a particular dwelling, indexed 'i', the 'averaged' GRP within a circular buffer centered on this dwelling was calculated as follows:

$$\text{'averaged' } GRP_i = \frac{\sum_k S_k * A_{i,k}}{\sum_k A_{i,k}}$$

in which:

- $A_{i,k}$ is the total area within the disc centered on dwelling 'i', in which GRP is classified in category 'k',

- S_k is, as defined above, the quantitative score allocated to areas in which GRP is classified in category 'k'.

Circular buffers centered on the 10,843 RC measurement points of radius 100 m, 200 m, 500 m, 1 km, 2 km, 5 km, 10 km, 20 km, 50 km and 100 km were used (Fig. 1.a).

The buffers were considered in order to investigate, without any a priori choice, which buffer was most associated with the RC (referred to as 'surrounding' GRP).

2.4 Statistical analysis

The observed statistical distribution of the measured radon concentrations being very close to log-normality (Fig. 2), all the statistical analyses were conducted using the decimal logarithms of radon concentration.

Multivariate statistical models were fitted, first, in order to identify the linear combination of buffers that provided the 'averaged' GRP most associated with logRC, then referred to as the 'surrounding' GRP, and, secondly, in order to assess the joint or specific effect of housing characteristics, residents' lifestyle, 'pinpoint' GRP and 'surrounding' GRP on logRC. The model chosen for this purpose was a

Author's copy

generalized least squares (GLS) multiple regression model that took into account the spatial correlation of logRC:

$logRC = \beta \, . \, X + E$

in which:

- X is a vector of covariables that may include GRP or 'surrounding' GRP, building characteristics and the lifestyle of the residents, and β is the vector of the regression coefficients,
- E is a vector with zero mean that models the part of logRC variability not explained by X. E includes instrumental errors as well as other characteristics not taken into account in X, including some aspects of geology, building characteristics, and residents' lifestyle. These characteristics may be micro-local, so that, as the distance d_{ij} between two dwellings 'i' and 'j' tends to zero, the correlation between the values $logRC_i$ and $logRC_j$ does not necessary tend to 1.

The spatial correlation between $logRC_i$ and $logRC_j$ was modeled using an exponential correlation structure:

$$\text{correlation}(logRC_i, logRC_j) = (1 - n).\exp\left(-\frac{d_{i,j}}{r}\right)$$

in which:

- d_{ij} is the distance between the measurement points 'i' and 'j',
- n, the 'nugget effect', is the fraction of the spatial variance of logRC that is related to spatially unstructured micro-local factors and therefore reduces the limit of the correlation between $logRC_i$ and $logRC_j$ to (1 – n) as the distance d_{ij} between the points 'i' and 'j' tends to zero,
- r, the 'range', is the distance for a reduction of the correlation between two measurements $logRC_i$ and $logRC_j$ by a factor $\exp(-1) \approx 0.37$.

GLS regressions were performed with the gls function of the 'nlme' package in R software[1].

[1]R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Author's copy

To account for the possible correlations between GRP and housing/lifestyle characteristics, the variations in the arithmetic means of logRC between the five GRP classes were examined for each of the various categories of individual housing/lifestyle characteristics.

3. Results

3.1. Indoor radon concentrations: general statistics

The RC varied from a few Bq/m^3 to 4,382 Bq/m^3 with a median of 49.4 Bq/m^3. The geometric mean RC was 53.0 Bq/m^3 with standard deviation of 2.55. In all, 22.8% of the concentrations were higher than 100 Bq/m^3 and 2.4% higher than 400 Bq/m^3. The spatial distribution of the measurement points covered the majority of France. However, for some areas, only a few results (or even no results) were available. Higher radon concentrations were generally observed in areas in which Hercynian granitic rocks are present, such as in western ('Massif Armoricain') and central ('Massif Central') France.

3.2. Association of radon concentration with housing and lifestyle characteristics

The geometric mean radon concentration was calculated by category of housing characteristics and lifestyle factors (Table 1). With regard to building materials, the highest geometric mean (171.5 Bq/m^3) was observed for the category consisting of 380 houses built with granite (Table 1). For the other materials, the means ranged from 36.5 Bq/m^3 for concrete to 67.1 Bq/m^3 for stone. With regard to building type, the lowest mean was reported for the 920 apartment buildings (30.9 Bq/m^3).The geometric mean concentration measured in the buildings constructed before 1945-1948 (65.8 Bq/m^3) was higher than that for more recent dwellings. For the other housing characteristics (foundations, type of room and floor level) the geometric mean radon concentrations was higher for the measurements made on the ground floor (compared to higher floors), and for the dwellings constructed on slabs or beaten earth. With regard to air renewal frequency, the means were relatively similar for the various classes and the mean was higher when there was no ventilation, as expected.

A GLS regression of logRC was fitted over the housing characteristics and residents' lifestyle data (see 2.4).The spatial model which fitted the data best was an exponential model with the following

Author's copy

estimated parameters: r = 114.8 km for the range and n = 0.60 for the nugget effect of the exponential correlation structure. The percentage of the total variation of logRC explained by the housing and lifestyle characteristics was 7.9% (Table 2). The coefficients of the GLS regression are shown in Table 1. The house-specific and lifestyle factors generally had a statistically significant association with logRC.

3.3. Association of radon concentration with geogenic radon potential

The GRP was highly variable and the areas covered by the various GRP classes were as follows: 53% of the area of France for class 1, 31% for class 2, 10% for class 3, 4.5% for class 4 and 1.5% for class 5.

The statistical distribution of RC by GRP class is shown in Table 3.The geometric mean RC was lowest for GRP classes 1 and 2 (46.9 and 48.1 Bq/m^3, respectively). The geometric mean RC increased from class 3 (82.0 Bq/m^3) to 5 (185.5 Bq/m^3). In GRP classes 1 and 2, the difference between the 80^{th} and 20^{th} percentiles of RC distribution was nearly 70 Bq/m^3 while in classes 3, 4 and 5, the differences were 140.4, 231.8 and 390.8 Bq/m^3, respectively. The percentage of RC above 100 Bq/m^3increased four-fold from the first (18%) to the last (69%) GRP class and the percentage of RC above 400 Bq/m^3 increased from 1% to 23%.

The scores (see 2.3) calculated for the 5 GRP classes (S_k, k=1;…; 5) were equal to 1.67, 1.68, 1.91, 2.13, and 2.27, respectively (Table 3).The percentage of the total variability of logRC explained by the 'pinpoint' GRP (GRP for the location of the dwelling in which a RC measurement was made), divided into five classes, was 4.9% (Table 2).

As explained in section 2.3, 'surrounding' GRP was then considered. First, an 'averaged' GRP was calculated over discs of various radii centered on each RC measurement point. The results showed that the correlations between logRC measured at a given point and the 'averaged' GRP over a disc centered on the point varied from 0.31 to 0.40. The highest correlations were observed for the discs of radius 20 km. The combination of 'averaged' GRP value over buffers that displayed the most significant multiple-correlation with logRC, was found to be the 2 km and 20 km buffers, which was

Author's copy

equivalent to the 'averaged' GRP calculated for the combination of the disc of radius 2 km and the ring around that disc and extending to the 20 km radius (Fig1.a and Fig 1.b.). This combination was therefore selected as the 'surrounding' GRP for a given point.

The GLS regression coefficients of logRC over the 'surrounding' GRP were 0.48 (95%CI: [0.37; 0.59]) for the 2 km disc; 0.90 (95% CI: [0.63; 1.16]) for the 2−20 km ring and −0.70 (95%CI: [−1.10; −0.29]) for the intercept. The model adjusted over the 'surrounding' GRP explained 14.4% of the total variability of logRC (Table 2).

When the 'surrounding' GRP range was subdivided into 12 classes of equal range (i.e. 0.05), the geometric mean of the RC increased from 43.1 Bq/m^3 (95% CI: [42.3; 43.9]) for the first class to 257.8 Bq/m^3 (95%CI: [195.1; 340.6]) for the last class (Fig. 3).The range between the 20th and 80th percentiles was a little wider for the last class but remained stable for the other classes. The RC values varied by a factor of about 4 from the 20th to the 80th percentiles in each of the classes. The percentage of RC above 100 Bq/m^3 increased from 15.4% to 78.4% as the 'surrounding' GRP increased from the first to last class (Fig.4).The percentage of RC above 400 Bq/m^3 increased from 0.8% to 35.1% as the 'surrounding' GRP increased from the first to last class (Fig.4).

3.4. Association of indoor radon concentration with geogenic radon potential and housing and lifestyle characteristics

Housing and lifestyle characteristics were, as expected, associated with local characteristics including geological characteristics. For example, houses built of granite were much more often found in areas where granite was present than in areas where sedimentary rocks were present. Regarding the possible correlations between GRP and housing/lifestyle characteristics, the results showed that for the categories of individual housing/lifestyle characteristics for which at least 10 RC values were available, the arithmetic mean of logRC was always found to increase from the first to the fifth GRP class (data not shown) within each class of house-specific/lifestyle factors.

The model including the 'surrounding' GRP and all the variables describing housing/lifestyle characteristics explained 20.6% of the total variation of logRC (Table 2). The 'surrounding' GRP was

Author's copy

associated with logRC with statistically significant estimated regression parameters. Some of the estimated regression coefficients of logRC over the housing/lifestyle characteristics were smaller (Table 4) than those obtained without considering the association of logRC with GRP (Table 1). However, those that were statistically significant when GRP was not considered remained statistically significant when GRP was taken into account.

4. Discussion

The aim of this paper was to describe the statistical associations between geology, house-specific/lifestyle characteristics and indoor radon concentration with a view to assisting the authorities in radon risk management, particularly by defining priority areas, and providing useful exposure data for epidemiological studies. This is the first French national study to analyze housing and geological factors influencing radon concentration over the entire country.

The geometric mean indoor radon concentrations estimated from the IRSN/Ministry of Health measurement survey was 53.0 Bq/m^3. The finding is consistent with those of previous French studies (Billon et al., 2005; Gambard et al., 2000; Kirchner et al. 2006). Compared to other European countries, indoor radon levels in France are similar to those in Belgium (Poffijn et al., 1985) and Spain (Quindos et al., 1993), higher than those in the United Kingdom (Green et al., 2002), and lower than those in Switzerland (Hauri et al., 2012), Hungary (Minda et al., 2009), Austria (Friedmann, 2005) and Finland (Mäkeläinen et al., 2009).

The housing parameters found to be the most strongly statistically associated with indoor radon concentration were the categories of building material, floor on which the measurement was made, foundation type, building type, and year of construction. The factors are those that have been frequently reported in the literature (Andersen et al., 2007; Gunby et al., 1993; Hauri et al., 2012; Hunter et al., 2009; Makelainen et al., 2001; Paillard et al., 2012; Papaefthymiou et al., 2003; Pirard et al., 2000; Sundal et al., 2004). In this study, lifestyle characteristics, such as air renewal frequency and ventilation, were not, or not strongly, associated with indoor radon concentration. These parameters are known to influence radon concentrations in dwellings. However, they were estimated by

Author's copy

qualitative values based on interview and were certainly not sufficiently precisely described. Overall, housing/lifestyle characteristics explained 7.9% of the total indoor radon concentration variation. Misclassifications in some answers to the questionnaire used to determine house-related and lifestyle factors may have contributed to the large percentage of unexplained variability.

This study also demonstrated that there was close agreement between geogenic radon potential and average indoor radon concentration: among the areas found to be more prone to high GRP, there were, on average, higher indoor radon concentrations and RC was more frequently greater than 100 Bq/m^3 or 400 Bq/m^3. When an 'averaged' GRP that took the surrounding geological information into account was calculated, the link between indoor radon concentration and geogenic radon potential became stronger. The 'averaged' GRP the most strongly associated with RC was found by considering the 'surrounding' GRP within an area combining the 2 km and 20 km discs centered on each RC measurement point. It is to be noted that those findings may depend strongly on the structuring of the spatial correlations of the indoor radon concentrations and may therefore not be extrapolated to other countries. The model showed that the percentage variability explained by the 'surrounding' GRP was 14.4% and was therefore greater than the percentage explained by the housing/lifestyle characteristics (7.9%). The results also showed that the 'surrounding' GRP was satisfactorily associated with average RC and the prevalence of RC to be greater than a specified threshold around a given point. It was thus possible to subdivide the total range of the 'surrounding' GRP values into 12 equal classes with markedly differing statistical distributions of exposure to radon. However, an individual assessment of radon exposure based on this model would be very limited since the RC variability within the 'surrounding' GRP classes remained high.

Earlier studies demonstrated the major influence of geology on indoor radon concentrations (Ielsch et al., 2002; Miles and Appleton, 2005; Minda et al., 2009; Sundal et al., 2004). In the model reported by Appleton and Miles (2010), 25% of the total variation in indoor radon concentrations in England and Wales was explained by bedrock and superficial geology. In the study reported herein, the GRP used was the GRP determined for each geological unit of the geological map of France on the 1:1,000,000 scale (Ielsch et al., 2010). The national geological map has an average positional uncertainty of 1km.

Author's copy

In consequence, misclassification may have occurred due to assignment of an incorrect GRP to indoor radon results. Actually, the 1:1, 000 000 scale is not the most appropriate for taking into account local geological variations and thus assessing GRP on the scale of an individual dwelling. If more detailed geological maps (e.g. 1:50,000 scale) and/or more detailed geological characteristics (e.g. information on local permeability) had been used, geology would probably have explained a larger proportion of the RC variation.

The best model, which included the 'surrounding' GRP and variables describing housing/lifestyle characteristics, explained about 20% of the overall variability of logRC. A number of factors may influence the fraction of the total variability of logRC explained by the housing characteristics: the lifestyle and geological covariates considered in a given study, including: the size and degree of geological heterogeneity of the area investigated (e.g. national or regional); the size and representativeness of the sample of measurements available; the quality of the time integrated RC measurements; the validity, precision and exhaustiveness of the information available on the geological, dwelling and lifestyle characteristics; and the validity and precision of the location of the dwellings and geological structures considered.

Compared to other predictive models developed, for example, in European countries, the results reported herein are similar to those obtained with the model developed in Switzerland (Hauri et al., 2012), in which 20% of the indoor radon concentration variability was explained by geological data and housing characteristics, and to those obtained with the model developed in England (Hunter et al., 2009), in which 29% of the variation was explained by geological units and housing parameters. In another study, which was conducted in Denmark (Andersen et al., 2007), 40% of the variability of almost 3,000 log-transformed indoor radon levels was explained by the model. That is more than in most other countries. In the Danish model, the most important explanatory variables were house type and geology. Geology explained a part of the variability that was similar to that explained by geology in the present study. However, the variability explained by housing characteristics was greater than that explained in the present study. This may be due to a lower ability of the housing characteristics recorded in the present study to take small-scale variations in radon levels into account. In order to

Author's copy

investigate the influence of geological and specific house type parameters on indoor radon, analyses of variance have been conducted in nationwide or large regional-scale studies. In the United Kingdom (Gunby et al., 1993) and Canada (Levesque et al., 1997), 21.7% of the variability of more than 2,000 indoor measurements and 15% of the variability of first floor measurements were explained. These results are similar to those obtained by Apte et al. (1999) who used a Bayesian mixed-effect linear regression model to predict indoor radon concentrations in New Hampshire (R^2=0.26). In Germany, analyses of variance of the data from three regions with 1,000 to 2,500 measurements each showed that the basement explained the major part of the indoor radon variability, of which about 36% remained unexplained (Kemski et al., 2009).

Several sources of variation may be included in the variability and spatial correlation of the residual variability such as an instrumental error component, the existence of other local characteristics of GRP, and dwelling and lifestyle characteristics, which are not exhaustively described by the information available for the present study. Some sources of variability may vary more or less continuously over space, while others may vary discontinuously, in particular, under the influence of properties of the 'micro-local environment'. The high value of the nugget effect (0.60 to 0.70 for various regression models) is also related to the limited information provided by the housing and lifestyle data available in this study. In Iowa, a geostatistical hierarchical Bayesian model based on 2,590 indoor radon concentration values was applied (Smith and Field, 2007) and accounted for errors in prediction due to known sources of variability. The authors showed that the measurement error was less important than unmeasured differences between homes and spatial dependencies in radon variability.

In the present study, since simple descriptive statistics may be biased due to spatial correlation, regressions were performed using a model that accounted for the spatial structure of logRC. Geostatistical methods would undoubtedly generate complementary results with regard to estimating radon exposure in France, as neighboring values of logRC may improve the prediction of logRC at a given point, in addition to the 'surrounding' GRP. A study is currently in progress at the IRSN with a

Author's copy

view to produce radon maps by testing different geostatistical models (kriging, co-kriging, kriging with external drift).

The results reported in the present paper demonstrate the major influence of GRP on the variability of the indoor radon concentrations at a regional or national scale. GRP could be of interest with regard to defining and mapping radon-prone areas in which the percentage of dwellings with high indoor radon concentrations may be the highest. So doing might be of assistance to the Authorities in radon risk management. Knowledge of geogenic radon potential could be particularly helpful in areas where no or only a few radon measurement results are available. The last model showed that taking GRP and housing/lifestyle factors into account may significantly improve indoor radon assessment. For epidemiological studies, using 'surrounding' GRP may be of more value than attempting to obtain indoor measurements, which may lead to selection biases due to very limited participation rates. For example, the results reported herein are being used in the GEOCAP research program (GEOlocalization of Pediatric CAncers, (Sermage-Faure et al., 2012)) to investigate the role of residential environmental exposures (including natural radioactivity) with regard to childhood leukemia in France. The potential impact of natural ionizing radiation on childhood leukemia has recently been studied in Denmark (Raaschou-Nielsen et al., 2008) and in Great Britain (Kendall et al., 2012), where predictive models were used to estimate the radon exposure of subjects with leukemia.

5. Conclusion

This study addressed the statistical association between indoor radon levels (data set of about 10,000 dwellings) and geogenic radon potential, house-specific factors and lifestyle characteristics. Geogenic radon potential was found to have the most significant influence on indoor radon concentrations. Housing/lifestyle characteristics explained only 7.9% of radon concentration variability. When geological information was added, 20% was explained. This finding is similar to the results reported with other recent predictive models in other countries. Consideration of the surrounding geogenic information was shown to increase the predictive capability of the model. The local unexplained variability of indoor radon concentration remains a limitation with respect to predicting radon

Author's copy

exposure at the individual level. However, the surrounding geogenic radon potential information was shown to assess satisfactorily a probable average radon concentration, and the prevalence of radon exposure was shown to be greater than a specified threshold on a multilevel scale. The present study also demonstrated that combining the two datasets, one based on indoor radon measurements and the other on geological features, enabled enhanced assessment of radon exposure in a given area in France. The results will be of assistance to programs designed to reduce or prevent high exposures to radon and will enable definition of the radon-prone areas in France. The results provide useful exposure data for epidemiological investigations on the health risk related to radon exposure, such as lung cancer and other cancers which have been studied more recently, such as childhood leukemia.

Author's copy

References

Andersen, C.E., Raaschou-Nielsen, O., Andersen, H.P., Lind, M., Gravesen, P., Thomsen, B.L., Ulbak, K., 2007. Prediction of 222Rn in Danish dwellings using geology and house construction information from central databases. Radiat Prot Dosimetry 123, 83−94.

Andru, J., 1998. Practical radon detection by Solid State Nuclear Track Detectors. Radioprotection 33, 501−513.

Appleton, J.D., Miles, J.C.H., 2010. A statistical evaluation of the geogenic controls on indoor radon concentrations and radon risk. Journal of Environmental Radioactivity 101, 799−803.

Apte, M.G., Price, P.N., Nero, A.V., Revzan, K.L., 1999. Predicting New Hampshire indoor radon concentrations from geologic information and other covariates. Environmental Geology 37, 181−194.

Baysson, H., Billon, S., Laurier, D., Rogel, A., Tirmarche, M., 2003. Seasonal correction factors for estimating radon exposure in dwellings in France. Radiat Prot Dosimetry 104, 245−252.

Billon, S., Morin, A., Caer, S., Baysson, H., Gambard, J.P., Backe, J.C., Rannou, A., Tirmarche, M., Laurier, D., 2005. French population exposure to radon, terrestrial gamma and cosmic rays. Radiat Prot Dosimetry 113, 314−320.

Friedmann, H., 2005. Final results of the Austrian Radon Project. Health Physics 89, 339−348.

Friedmann, H., Groller, J., 2010. An approach to improve the Austrian Radon Potential Map by Bayesian statistics. Journal of Environmental Radioactivity 101, 804−808.

Gambard, J.P., Mitton, N., Pirard, P., 2000. Campagne nationale de mesure de l'exposition domestique au radon IPSN-DGS. Bilan et représentation cartographique des mesures au 01 janvier 2000. Institut de Protection et de Sûreté Nucléaire

Green, B., Miles, J., Bradley, E., Rees, D., 2002. NRCP- W26 Radon atlas of England and Wales.

Gunby, J.A., Darby, S.C., Miles, J.C.H., Green, B.M.R., Cox, D.R., 1993. Factors affecting indoor radon concentrations in the United-Kingdom. Health Physics 64, 2−12.

Hauri, D.D., Huss, A., Zimmermann, F., Kuehni, C.E., Röösli, M., 2012. A prediction model for assessing residential radon concentration in Switzerland. Journal of Environmental Radioactivity 112, 83-89.

Hunter, N., Muirhead, C.R., Miles, J.C.H., Appleton, J.D., 2009. Uncertainties in radon related to house-specific factors and proximity to geological boundaries in England. Radiation Protection Dosimetry 136, 17−22.

Ielsch, G., Cushing, M.E., Combes, P., Cuney, M., 2010. Mapping of the geogenic radon potential in France to improve radon risk management: methodology and first application to region Bourgogne. J Environ Radioact 101, 813−820.

Ielsch, G. and Cuney, M., 2004. Cartographie prédictive du potentiel d'exhalation du radon-222 à la surface des sols : exemple d'application dans le Massif Armoricain. Environnement, Risques et Santé 2004. Vol. 3, n°1, janvier-février.

Ielsch, G., Ferry, C., Tymen, G., Robe, M.C., 2002. Study of a predictive methodology for quantification and mapping of the radon-222 exhalation rate. Journal of Environmental Radioactivity 63, 15−33.

Author's copy

Ielsch, G., Thieblemont, D., Labed, V., Richon, P., Tymen, G., Ferry, C., Robe, M.C., Baubron, J.C., Bechennec, F., 2001. Radon (Rn-222) level variations on a regional scale: influence of the basement trace element (U, Th) geochemistry on radon exhalation rates. Journal of Environmental Radioactivity 53, 75−90.

Kemski, J., Klingel, R., Siehl, A., Valdivia-Manchego, M., 2009. From radon hazard to risk prediction-based on geological maps, soil gas and indoor measurements in Germany. Environmental Geology 56, 1269−1279.

Kendall, G.M., Little, M.P., Wakeford, R., Bunch, K.J., Miles, J.C., Vincent, T.J., Meara, J.R., Murphy, M.F., 2012. A record-based case-control study of natural background radiation and the incidence of childhood leukaemia and other cancers in Great Britain during 1980-2006. Leukemia.

Kirchner, S., Arenes, J.F., Cochet, C., Derbez, M., Duboudin, C., Elias, P., Gregoire, A. , Jédor, B., Lucas, J.P., Pasquier, N., Pigneret, M., Ramalho, O., 2006. Campagne nationale logements : Etat de la qualité de l'air dans les logements français. Rapport final. Observatoire de la Qualité de l'Air Intérieur, DDD/SB 2006-57.

Levesque, B., Gauvin, D., McGregor, R.G., Martel, R., Gingras, S., Dontigny, A., Walker, W.B., Lajoie, P., Letourneau, E., 1997. Radon in residences: Influences of geological and housing characteristics. Health Physics 72, 907−914.

Makelainen, I., Arvela, H., Voutilainen, A., 2001. Correlations between radon concentration and indoor gamma dose rate, soil permeability and dwelling substructure and ventilation. Science of the Total Environment 272, 283−289.

Mäkeläinen, I., Kinnunen, T., Reisbacka, H., Valmari, T., Arvela, H., 2009. Radon in Finnish dwellings - Sample survey 2006. Radiation and Nuclear Safety Authority, Helsinki.

Miles, J.C.H., Appleton, J.D., 2005. Mapping variation in radon potential both between and within geological units. Journal of Radiological Protection 25, 257−276.

Minda, M., Toth, G., Horvath, I., Barnet, I., Hamori, K., Toth, E., 2009. Indoor radon mapping and its relation to geology in Hungary. Environmental Geology 57, 601−609.

Paillard, J., Roudier, C., Lucas, J., Zeghnoun, A., Catelinois, O., Pirard, P., 2012. Factors influencing indoor radon levels in French dwellings. ISEE 2012 conference, Columbia, South Carolina.

Papaefthymiou, H., Mavroudis, A., Kritidis, P., 2003. Indoor radon levels and influencing factors in houses of Patras, Greece. Journal of Environmental Radioactivity 66, 247−260.

Pirard, P., Richardson, S., Laurier, D., Mitton, N., Gambard, J., Monfort, C., 2000. Analyse écologique de la relation entre l'exposition au radon dans l'habitat et la mortalité par cancer du poumon en France. Rapport final dans le cadre du Programme Environnement-Santé 1996, Convention EN96C5 et EN96D5. Institut de Protection et de Sûreté Nucléaire, SEGR/LEADS/NT/2000-23.

Poffijn, A., Marijns, R., Vanmarcke, H., Uyttenhove, J., 1985. Results of a preliminary survey on radon in Belgium. Science of the Total Environment 45, 335−342.

Quindos, L.S., Fernandez, P.L., Soto, J., 1993. Exposure to natural sources of radiation in Spain. Nuclear Tracks and Radiation Measurements 21, 295−298.

Author's copy

Raaschou-Nielsen, O., Andersen, C.E., Andersen, H.P., Gravesen, P., Lind, M., Schuz, J., Ulbak, K., 2008. Domestic radon and childhood cancer in Denmark. Epidemiology 19, 536−543.

Sermage-Faure, C., Laurier, D., Goujon-Bellec, S., Chartier, M., Guyot-Goubin, A., Rudant, J., Hemon, D., Clavel, J., 2012. Childhood leukemia around French nuclear power plants--the Geocap study, 2002-2007. International journal of cancer. Journal international du cancer 131, E769−780.

Smith, B.J., Field, R.W., 2007. Effect of housing factors and surficial uranium on the spatial prediction of residential radon in Iowa. Environmetrics 18, 481−497.

Sundal, A.V., Henriksen, H., Soldal, O., Strand, T., 2004. The influence of geological factors on indoor radon concentrations in Norway. Science of the Total Environment 328, 41−53.

Verger, P., Hubert, P., Cheron, S., Bonnefous, S., Bottard, S., Brenot, J., 1994. Use of field-measurements in radon mapping in France. Radiation Protection Dosimetry 56, 225−229.

Zhu, H.C., Charlet, J.M., Poffijn, A., 2001. Radon risk mapping in southern Belgium: an application of geostatistical and GIS techniques. Science of the Total Environment 272, 203−210.

Zhu, H.C., Charlet, J.M., Tondeur, F., 1998. Geological controls to the indoor radon distribution in southern Belgium. Science of the Total Environment 220, 195−214.

Author's copy

Fig.1.a.Quantification of the surrounding geogenic radon potential: in this example, the average value of GRP ('averaged' GRP) is determined in a circular buffer, with a radius of 20 km, centered on dwelling 'i'. The disc includes 5 zones: $A_{i,k}$ is the area of the zone in which GRP is classified in category 'k' (see text section 2.3).

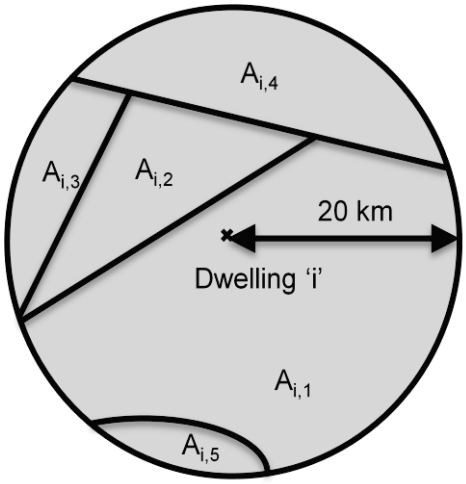
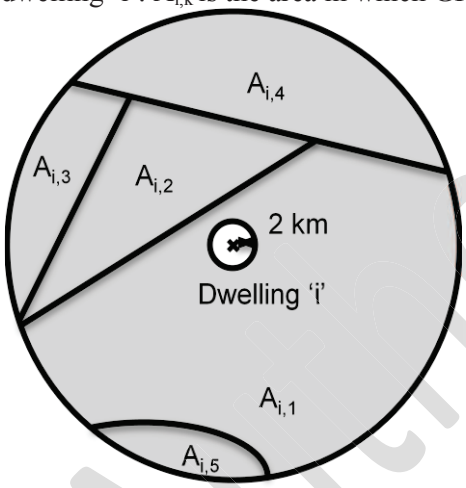
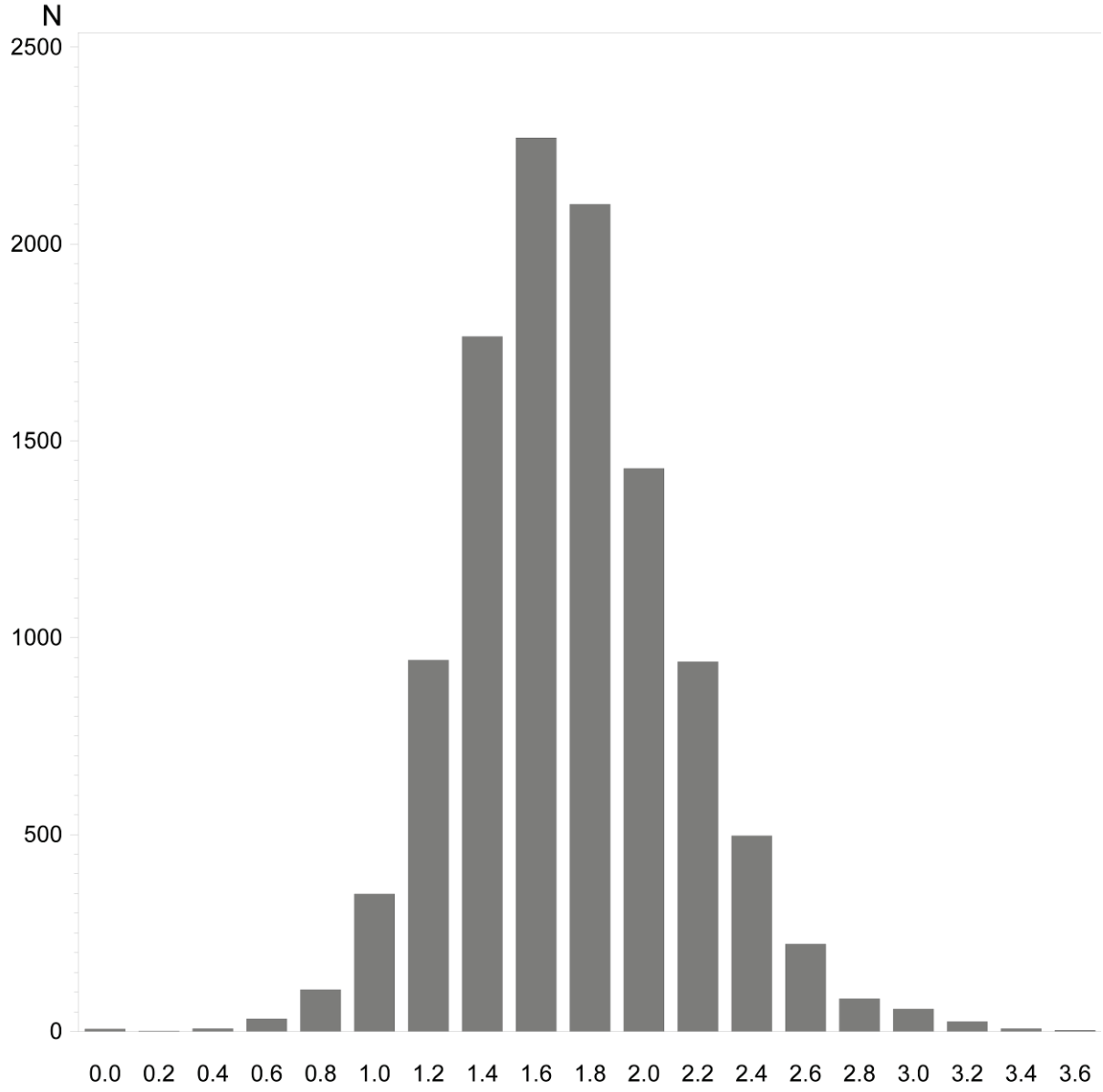


Fig.1.b. Quantification of the surrounding geogenic radon potential: in this example, the average value of GRP ('averaged' GRP) is determined over a 2 to 20 km ring (colored in grey) centered on the dwelling 'i'. $A_{i,k}$ is the area in which GRP is classified in category 'k' (see section 2.3 for details).



Author's copy

Fig.2. Histogram of the decimal logarithms (logRC) of the 10,843 results of the indoor radon concentration measurements.

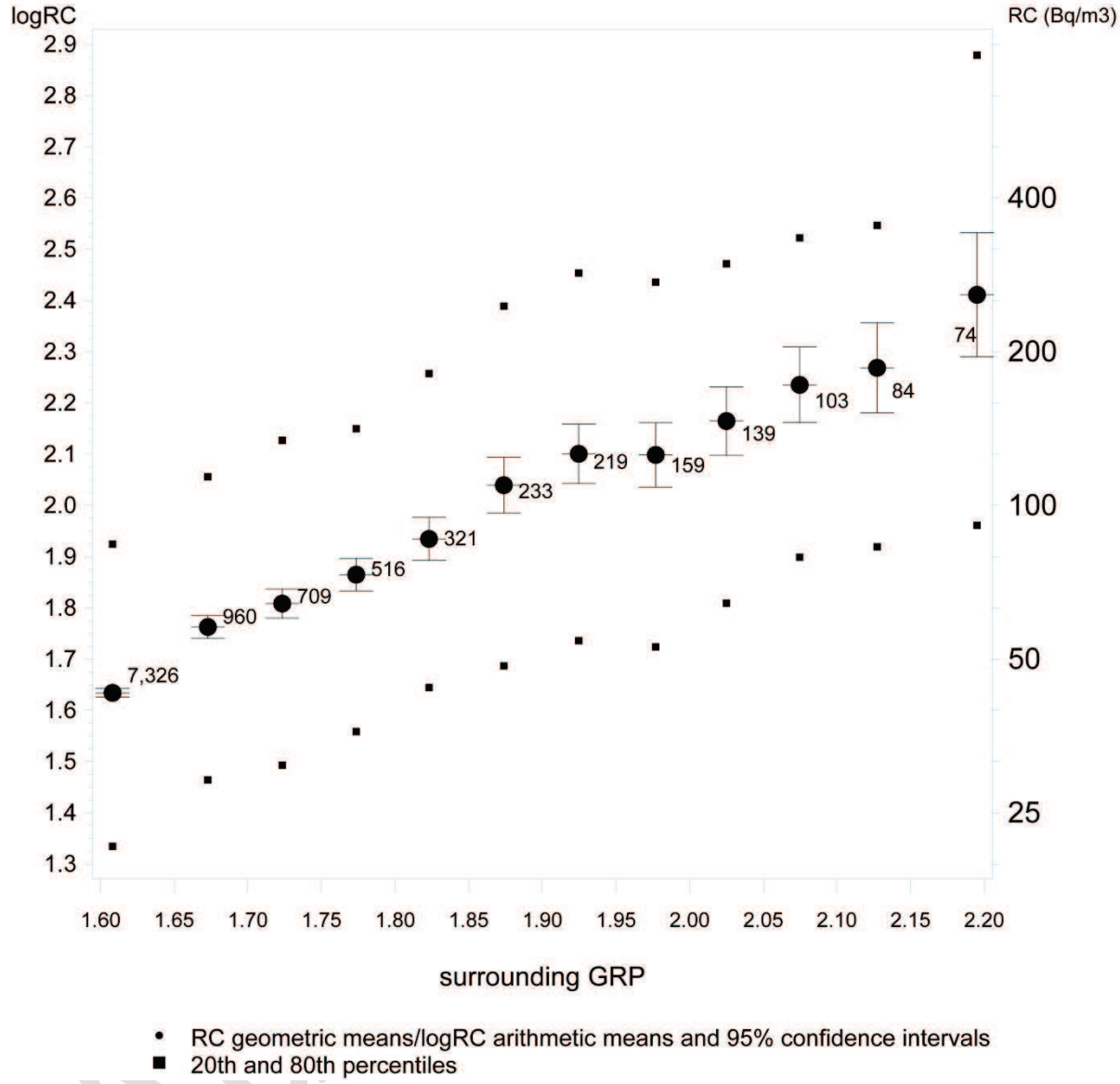


RC: radon concentration measured in each dwelling

Abscissa: values of logRC; using this scale, logRC of 0, 1, 2 or 3 is equivalent to RC of 1, 10, 100 or 1000 Bq/m^3

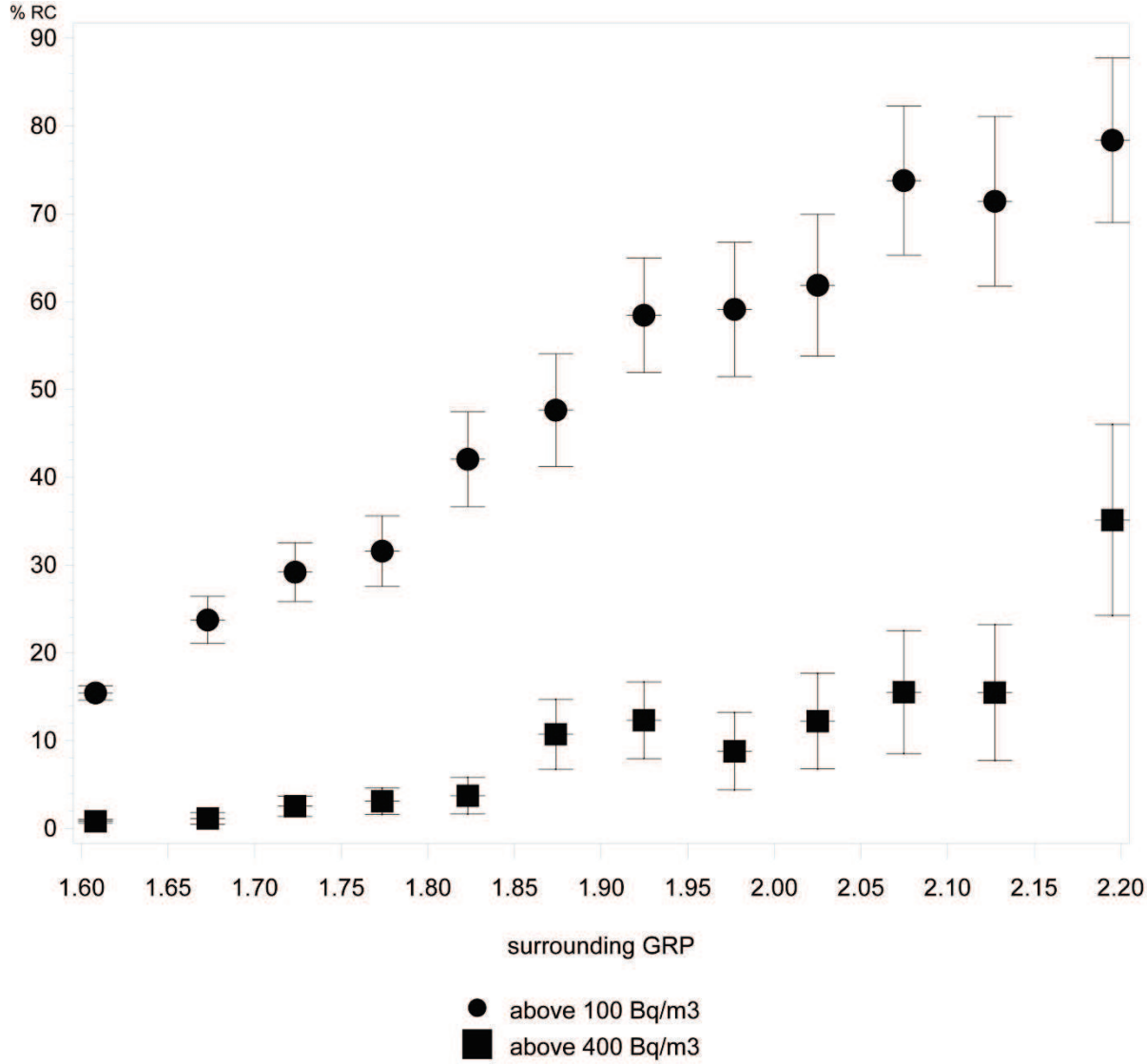
Ordinate: N = number of results

Fig.3. Statistics of the logarithm of the indoor radon concentration (logRC) and number of radon concentration measurements by category of surrounding geogenic radon potential (GRP).The total range of the ‘surrounding’ GRP was divided into 12 equal classes.



Abscissa: arithmetic means of the ‘surrounding’ GRP in each class.

Fig.4. Observed proportions of RC above 100 Bq/m^3 and 400 Bq/m^3 by category of surrounding geogenic radon potential (GRP) and their 95% confidence intervals. The total range of the 'surrounding' GRP was divided into 12 equal classes.



Abscissa: 'surrounding' GRP = arithmetic means of the 'surrounding' GRP in each class.

Ordinate: %RC = Observed proportion of RC above 100 Bq/m^3 and 400 Bq/m^3 by category of surrounding geogenic radon potential (GRP) and their 95% confidence intervals

Table 1: Geometric mean radon concentrations (Bq/m^3) and estimated regression coefficients[a] of the logarithm of radon concentrations (logRC) for each housing characteristic and lifestyle factor

		n [b]	Radon concentrations: geometric mean [95% CI] [c]	Multivariate GLS regression 10^{β} [95% CI] [d]	p-value [e]
	Intercept	-	42.45 [30.49; 59.11]		***
Date of construction	After 1945-1948	5,669	44.4 [43.7; 45.7]	0.82 [0.78; 0.86]	***
	Before 1945-1948	4,855	65.8 [64.6; 67.6]	**1.0 [ref.]**	-
	Unknown	319	46.7 [42.7; 51.3]	-	
Building material	Wood	193	45.8 [39.8; 52.5]	0.92 [0.81; 1.06]	ns
	Concrete	850	36.5 [34.7; 38.9]	1.05 [0.99; 1.05]	ns
	Breeze-block	2,667	49.2 [47.9; 51.3]	1.06 [1.01; 1.12]	*
	Granite	380	171.5 [154.9; 190.6]	1.76 [1.58; 1.96]	***
	Stone	3,447	67.1 [64.6; 69.2]	1.16 [1.09; 1.23]	***
	Other	289	56.7 [51.3; 63.1]	1.14 [1.02; 1.26]	*
	Brick	2,489	40.8 [39.8; 42.7]	**1.0 [ref.]**	-
	Unknown	528	45.8 [42.7; 49.0]	-	
Foundations	Basement or cellar	5,592	48.6 [47.9; 50.1]	0.85 [0.82; 0.88]	***
	Slab or beaten earth	3,677	62.4 [60.3; 64.6]	**1.00 [ref.]**	-
	Unknown	1,574	49.4 [46.8; 51.3]	-	
Building type	House	9,194	57.0 [56.2; 57.5]	1.41 [1.31; 1.52]	***
	Other	562	41.6 [38.9; 44.7]	1.20 [1.06; 1.35]	**
	Apartment building	920	30.9 [29.5; 32.4]	**1.0 [ref.]**	-
	Unknown	167	45.0 [38.9; 51.3]	-	
Room type	Bedroom	1,113	50.4 [47.9; 53.7]	1.02 [0.96; 1.08]	ns
	Other	1,066	46.5 [43.7; 49.0]	1.07 [0.99; 1.15]	†
	Living room or kitchen	8,559	54.4 [53.7; 55.0]	**1.0 [ref.]**	-
	Unknown	105	46.3 [38.0; 56.2]	-	
Floor level	Above the ground floor	2,228	43.2 [41.7; 44.7]	0.77 [0.74; 0.81]	***
	Ground floor	8,394	56.1 [55.0; 57.5]	**1.0 [ref.]**	-
	Unknown	221	50.8 [44.7; 57.5]	-	
Air renewal frequency	High	3,318	51.8 [50.1; 53.7]	1.01 [0.92; 1.12]	ns
	Medium	6,731	54.6 [53.7; 56.2]	1.07 [0.97; 1.18]	ns
	Low	392	51.7 [46.8; 57.5]	**1.0 [ref.]**	-
	Unknown	402	41.2 [38.0; 44.7]	-	
Ventilation	Natural	3,646	48.7 [46.8; 50.1]	0.96 [0.92; 1.00]	*
	Mechanical and individual	1,358	52.0 [50.1; 55.0]	0.97 [0.92; 1.03]	ns
	Mechanical and collective	212	33.8 [30.9; 37.2]	1.00 [0.89; 1.12]	ns
	None	5,247	58.2 [56.2; 60.3]	**1.0 [ref.]**	-
	Unknown	380	46.4 [41.7; 51.3]	-	

[a] Multivariate generalized least squares (GLS) regression of logRC for housing characteristics and lifestyle factors

[b] n = number of dwellings in which the radon concentration was measured

[c] [95%CI] = 95% confidence interval

[d] 10^{β} [95% CI] = ten to the power of the estimated GLS regression coefficient of logRC= multivariate adjusted geometric mean of RC for France for the intercept and ratio of the multivariate adjusted geometric mean RC in a category to the multivariate adjusted geometric mean RC in the reference category [ref.]

[e] significance level of the difference between the estimated ß and zero
***: p <0.001; ** : p <0.01; * : p <0.05; † : p<0.10; ns: p>0.10

Author copy

Table 2: Characteristics of the generalized least squares regressions over the 10,843 indoor radon concentration (RC) measurements

Model	Residual variance	Explained variance (%)	Exponential correlation parameters	
			Nugget [a] (n)	Range (km)[b] (r)
Model without any covariate	0.176	-	0.60	90.6
'Pinpoint'[c] GRP [d]	0.167	4.9%	0.60	98.3
Housing and lifestyle characteristics	0.162	7.9%	0.60	114.8
'Pinpoint' GRP, and housing and lifestyle characteristics	0.155	12.0%	0.62	126.4
'Surrounding' GRP [e]	0.150	14.4%	0.70	82.7
'Surrounding' GRP, and housing and lifestyle characteristics	0.140	20.6%	0.70	112.3

[a] n, the 'nugget effect' is the fraction of the spatial variance of the logarithm of the radon concentrations (logRC) that is related to spatially unstructured micro-local factors (see text § 2.4)

[b] r, the 'range', is the distance for a reduction in the correlation between two measurements $logRC_i$ and $logRC_j$ by a factor exp (-1) ≈ 0.37.

[c] The 'pinpoint' GRP is the GRP for the location of the dwelling where RC was measured (see text § 2.3).

[d] GRP = geogenic radon potential.

[e] The 'surrounding' GRP is the 'averaged' GRP calculated over a buffer with 2 and 20 km radii centered on the location of the dwelling where RC was measured (see text § 2.3).

Author's copy

Table 3: Distribution of the 10,843 indoor radon concentrations (RC) with respect to geogenic radon potential (GRP).

'Pinpoint' GRP Classes (k)	Number of RC measurements	'Pinpoint' GRP score S_k: arithmetic mean of logRC (standard deviation)	RC geometric mean [a] (Bq/m^3) (standard deviation)	RC P20% [b] – P80% [c] (Bq/m^3)	RC (%)	
					≥ 100 Bq/m^3	≥ 400 Bq/m^3
1	5,352	1.67 (0.37)	46.9 (2.33)	23.8 − 93.2	18%	1%
2	3,887	1.68 (0.40)	48.1 (2.48)	23.1 − 97.2	19%	2%
3	1,010	1.91 (0.42)	82.0 (2.64)	36.1 − 176.5	40%	6%
4	450	2.13 (0.43)	134.2 (2.68)	56.8 − 288.7	60%	11%
5	144	2.27 (0.46)	185.5 (2.88)	72.8 − 463.6	69%	23%
Total	10,843	1.72 (0.41)	53.0 (2.55)	25.0 − 110.8	23%	2%

[a] Geometric mean of the distribution of the RC
[b] 20th percentile of the distribution of the RC
[c] 80th percentile of the distribution of the RC

Author's copy

Table 4: Estimated regression coefficients of the logarithm of the radon concentrations (logRC) on the surrounding geogenic radon potential (GRP) and the housing/lifestyle characteristics considered simultaneously [a]

		10^{β} [95% CI] [b]	p-value [c]
	Intercept	0.29 [0.11;0.75]	*
'Surrounding' GRP	2 km	2.67 [2.05; 3.49]	***
	2-20 km ring	6.49 [3.52; 11.96]	***
Date of construction	After 1945-1948	0.82 [0.78; 0.86]	***
	Before 1945-1948	**1.0 [ref.]**	-
Building material	Wood	0.91 [0.80; 1.04]	ns
	Concrete	1.05 [0.97; 1.13]	ns
	Breeze-block	1.06 [1.00; 1.11]	*
	Granite	1.55 [1.39; 1.72]	***
	Stone	1.16 [1.09; 1.22]	***
	Other	1.14 [1.03; 1.26]	*
	Brick	**1.0 [ref.]**	-
Foundations	Basement or cellar	0.85 [0.82; 0.88]	***
	Slab or beaten earth	**1.0 [ref.]**	-
Building type	House	1.41 [1.31; 1.52]	***
	Other	1.21 [1.07; 1.36]	**
	Apartment building	**1.0 [ref.]**	-
Room type	Bedroom	1.02 [0.96; 1.08]	ns
	Other	1.07 [1.00; 1.15]	*
	Living room or kitchen	**1.0 [ref.]**	-
Floor level	Above the ground floor	0.77 [0.74; 0.81]	***
	Ground floor	**1.0 [ref.]**	-
Air renewal frequency	High	1.01 [0.92; 1.12]	ns
	Medium	1.08 [0.98; 1.18]	ns
	Low	**1.0 [ref.]**	-
Ventilation	Natural	0.96 [0.92; 0.99]	*
	Mechanical and individual	0.97 [0.92; 1.02]	ns
	Mechanical and collective	0.99 [0.88; 1.12]	ns
	None	**1.0 [ref.]**	-

[a] multivariate generalized least squares (GLS) regression of logRC on the surrounding geogenic radon potential and each housing/lifestyle characteristic

[b] 10^{β} [95% CI] = ten to the power of the estimated GLS regression coefficient of logRC on the different categories with the specified category [ref.] as reference category: the percentage increase in the geometric mean with respect to the reference category.

[c] significance level ***: $p<0.001$; ** : $p<0.01$; * : $p<0.05$; † : $p<0.10$; ns: $p>0.10$

Author's copy