

CANCER

DÉCEMBRE 2024

DONNÉES DE SURVEILLANCE

**OBSERVATOIRE DES CANCERS**  
**DU REIN, DE LA VESSIE**  
**ET DES LEUCÉMIES AIGUËS**  
**CHEZ L'ADULTE DANS LE DÉPARTEMENT**  
**DES BOUCHES-DU-RHÔNE (REVELA13)**

Analyses spatiales, 2013-2018

RÉGION

PROVENCE-ALPES-CÔTE D'AZUR

En partenariat avec :



**REVELA13**  
Observatoire

Observatoire des cancers du **RE**in, de la **VE**ssie et des **LE**ucémies **A**iguës  
chez l'adulte dans le département des Bouches-du-Rhône.

## Résumé

### Observatoire des cancers du rein, de la vessie et des leucémies aiguës chez l'adulte dans le département des Bouches-du-Rhône (REVELA13)

Analyses spatiales, 2013-2018

L'Observatoire REVELA13 enregistre, depuis 2012, les cancers incidents du REin, de la VEssie et les Leucémies Aiguës de l'adulte sur le département des Bouches-du-Rhône. L'enjeu de ce dispositif est la surveillance territorialisée de l'incidence de cancers ciblés et la détection d'éventuels agrégats spatio-temporels et/ou de zones de concentration de cancers.

Des premières analyses spatiales sur les données 2013-2016 ont été publiées en 2019. Ce second rapport présente les analyses complémentaires réalisées sur deux années supplémentaires de données.

Les ratios d'incidence standardisés (SIR) sur les données de population française ont été lissés par un modèle spatial bayésien et ajustés sur des facteurs connus pour influencer la répartition des cancers dans la population : inégalités sociales, accès aux soins, densité de population et sur un *proxy* du tabagisme (taux communal normalisé de mortalité par cancer du poumon), facteur de risque reconnu de ces cancers. La détection de regroupements de cas a été réalisée par des méthodes globales et une méthode locale de détection de *clusters* (agrégats spatio-temporels). Lorsque cela était possible, les analyses ont été conduites à deux échelles géographiques distinctes : la commune et l'Iris (Îlots regroupés pour l'information statistique).

Au niveau départemental, l'incidence estimée du cancer de la vessie était significativement supérieure à l'incidence nationale pour les deux sexes. Des surincidences de cancers de la vessie ont été observées chez les hommes et chez les femmes dans de nombreuses communes. Chez les hommes, le risque de cancer était associé positivement à la densité de population et au *proxy* du tabagisme. Deux clusters ont été détectés dans la commune : un était situé dans la partie sud-est du département, le second recouvrait l'agglomération de Marseille. À l'échelle de l'Iris, aucune association avec les facteurs d'ajustement n'était observée. Les Iris présentant une surincidence étaient recentrés sur Marseille et la partie sud-est du département. À cette échelle, la détection de clusters identifiait des regroupements d'Iris issus des clusters observés à l'échelle communale mais ne permettait pas de détecter d'autres clusters dans le département. Chez les femmes, aucun lien n'était mis en évidence entre le risque de cancer de vessie et les différents facteurs d'ajustement. Une surincidence significative a été observée pour la plupart des arrondissements de la ville de Marseille et une commune située au nord. La recherche de clusters a identifié également un agrégat de communes présentant un excès de risque sur une zone à peu près identique.

Les incidences départementales estimées du cancer du rein et des leucémies aiguës myéloïdes (LAM) n'étaient pas différentes des incidences nationales pour les deux sexes pendant la période étudiée. Les taux d'incidence du cancer du rein étaient répartis de façon hétérogène sur le département mais sans regroupements de cas identifiables. Aucune commune ne présentait de surincidence de cancer du rein et aucun cluster n'était détecté chez les hommes comme chez les femmes. Concernant les LAM, les cas étaient répartis de façon globalement homogène sur le territoire, cependant la cartographie des SIR lissés a mis en évidence un arrondissement de Marseille avec une surincidence significative uniquement chez les hommes. Aucun cluster n'était détecté dans la commune pour les deux sexes.

Compte tenu d'un grand nombre d'Iris sans cas observé, les analyses à cette échelle n'ont pas pu être menées pour les cancers de vessie chez la femme, les cancers du rein et les LAM.

Les résultats confirment l'existence d'une surincidence de cancers de vessie à l'échelle du département, plus marquée dans les communes densément peuplées et pour lesquelles un tabagisme important était estimé. Ces résultats suggèrent également l'existence d'autres facteurs explicatifs communs à ces zones. L'exposition des populations concernées à la pollution atmosphérique, facteur fortement suspecté d'être lié à ce cancer dans la littérature, reste ainsi une hypothèse à explorer. L'ajout de deux années supplémentaires de données a permis de mettre en évidence des zones d'intérêt pour les cancers du rein et les LAM.

**MOTS-CLÉS :** CANCER DE LA VESSIE, CANCER DU REIN, LEUCÉMIES AIGUËS, ANALYSES SPATIALES, DÉTECTION DE CLUSTERS

**Citation suggérée :** Mantey K, Lehot L, Gaudart J, Pascal L. Observatoire des cancers du rein, de la vessie et des leucémies aiguës chez l'adulte dans le département des Bouches-du-Rhône (REVELA13). Analyses spatiales, 2013-2018. Saint-Maurice : Santé publique France, 2024. 102 p.  
Disponible à partir de l'URL : [www.santepubliquefrance.fr](http://www.santepubliquefrance.fr)

ISSN : 2534-6539 ; ISBN-NET : 979-10-289-0931-4 - RÉALISÉ PAR LA DIRECTION DE LA COMMUNICATION, SANTÉ PUBLIQUE FRANCE - DÉPÔT LÉGAL : DÉCEMBRE 2024

## Abstract

### Observatory of kidney cancers, bladder cancer and acute leukemia in adults in the Bouches-du-Rhône department (REVELA13) Spatial analyses, 2013-2018

Since 2012, the REVELA13 Observatory records incident cases of kidney and bladder cancers and acute leukemia in the adult population of the Bouches-du-Rhône department. Through this system, the aim is to survey the geographical repartition of these cancers at a fine sub-departmental scale and identify potential clusters of cases across the department. Correlation studies with environmental factors can then be conducted. A first analysis of the data from 2013 to 2016 was published in 2019. This second report presents an update of the results with an extension of the study period up to 2018.

Incidence ratios were standardized (SIR) using French population and incidence data. SIR were then smoothed using a spatial Bayesian model while adjusting on factors known to influence cancer distribution in the population: social inequalities, access to healthcare, population density and a surrogate of smoking prevalence (standardized mortality ratio of lung cancer) – a well-known risk factors of cancers in this study. Cluster detection was realized through global indexes and a local method. Whenever possible, analysis were conducted at two different spatial scales: communes and Iris.

At a departmental scale, the estimation of the incidence of bladder cancer was significantly higher than the national incidence in both sexes. A significant excess of cases of bladder cancer was observed in many communes among both men and women. Among men, the risk of cancer was positively associated with population density and smoking *proxy*. Two clusters were detected at the commune level: one in the south-eastern part of the department and the second covering Marseille agglomeration. At the Iris level, no association with the adjustment factors was observed. Iris displaying a high incidence were gathered in Marseille and the south-eastern part of the department. At this scale, cluster detection identified groupings of Iris from clusters observed at the communal level but did not indentify any other clusters across the department. For women, no link was found between bladder cancer risk and the various adjustment factors. A significant high incidence was observed for almost all districts of the city of Marseille and one commune located in the North. The cluster search also identified an aggregate of communes with excess cases in an almost identical area.

The estimated departmental incidences of kidney cancer and acute myeloid leukemia (AML) were not different from the national incidences for both sexes over the study period. The incidence ratios of kidney cancer were heterogeneously distributed over the department but without any identifiable clusters of cases. No commune presented a significant high incidence of kidney cancer and no clusters were detected in either men or women. AML cases were globally homogeneously distributed over the region, however the mapping of smoothed SIRs highlighted a district of Marseille with a significant high incidence in men. No cluster was detected in the commune for either sex. Given the large number of Iris without any observed cases, analysis at the Iris level could not be carried out for bladder cancer in women, kidney cancer and AML.

These results confirm the existence of an excess of bladder cancer at the departmental level, even more in densely populated communes for which a high smoking prevalence was observed. These results also suggest the existence of other explanatory factors common to these areas. The exposure of the populations concerned to air pollution – a factor strongly suspected of being linked to this cancer in the literature, remains a hypothesis to be explored. The addition of two more years of data has made it possible to highlight areas of interest for kidney cancer and AML.

**KEY WORDS:** BLADDER CANCER, KIDNEY CANCER, ACUTE LEUKEMIA,  
SPATIAL ANALYSIS, CLUSTER DETECTION

## Auteurs

- **Karine Mantey**, Observatoire REVELA13 – Santé publique France en Provence-Alpes-Côte d’Azur et en Corse, Marseille, France
- **Laurent Lehot**, Observatoire REVELA13 – Aix Marseille Université, IRD, Inserm, Isspam, Sesstim, Marseille, France
- **Jean Gaudart**, Aix Marseille Université, IRD, Inserm, Isspam, Sesstim, AP-HM, Hôpital La Timone, BioSTIC, Biostistic & ICT unit, Marseille, France
- **Laurence Pascal**, Santé publique France en Provence-Alpes-Côte d’Azur et en Corse, Marseille, France

## Relecture

- **Érik-André Sauleau**, Laboratoire des sciences de l’ingénieur, de l’informatique et de l’imagerie (ICube), UMR7357, Université de Strasbourg

## Remerciements

Pour leur participation au recueil des données :

- les responsables des départements d’informations médicales, des services de chirurgie urologique et générale, d’hématologie et de médecine interne des établissements publics et privés et des centres de lutte contre le cancer ;
- les responsables des laboratoires d’anatomopathologie ;
- les réseaux régionaux en cancérologie OncoPACA-Corse et Onco Occitanie ;
- les centres de coordination en cancérologie des Bouches-du-Rhône, du Gard et du Vaucluse et plus particulièrement de l’AP-HM et de l’IPC.

Pour leur aide dans l’animation des réseaux :

- **Professeur Gilles Karsenty**, service d’urologie et de transplantation rénale de l’hôpital de la Conception à l’AP-HM
- **Professeur Norbert Vey**, responsable de l’unité de traitement des leucémies et de l’unité d’évaluation thérapeutique en onco-hématologie dédiée aux essais précoces à l’Institut Paoli-Calmettes
- **Professeur François Eisinger**, département d’anticipation et de suivi du cancer à l’Institut Paoli-Calmettes

# Sommaire

<b>Résumé</b> .....	<b>2</b>
<b>Abstract</b> .....	<b>4</b>
<b>Auteurs</b> .....	<b>5</b>
<b>Relecture</b> .....	<b>5</b>
<b>Remerciements</b> .....	<b>5</b>
<b>Abréviations</b> .....	<b>8</b>
<b>1 CONTEXTE ET OBJECTIFS</b> .....	<b>10</b>
1.1 <b>Contexte local</b> .....	<b>10</b>
1.2 <b>Objectifs</b> .....	<b>11</b>
<b>2 MATÉRIEL ET MÉTHODES</b> .....	<b>12</b>
2.1 <b>Période et zone d'étude</b> .....	<b>12</b>
2.2 <b>Populations et définitions de cas</b> .....	<b>12</b>
2.2.1 Recensement des cas.....	12
2.2.2 Gestion des données nominatives.....	13
2.2.3 Définitions de cas.....	14
2.2.4 Données de population .....	14
2.3 <b>Données géographiques</b> .....	<b>15</b>
2.3.1 Sources des données.....	15
2.3.2 Regroupement des Iris peu ou pas peuplés .....	15
2.3.3 Identification des communes et Iris de résidence des cas .....	16
2.4 <b>Facteurs d'ajustement</b> .....	<b>16</b>
2.4.1 Densité de population.....	16
2.4.2 Indices de défaveur sociale .....	17
2.4.3 Accès aux soins de premier recours.....	17
2.4.4 Tabagisme .....	18
2.5 <b>Analyse descriptive et estimation d'incidence</b> .....	<b>18</b>
2.6 <b>Cartographie des risques lissés</b> .....	<b>19</b>
2.7 <b>Recherche de clusters</b> .....	<b>21</b>
2.7.1 Méthodes globales de clustering .....	21
2.7.2 Méthode de détection locale non ciblée .....	22
<b>3 DESCRIPTION DE LA ZONE D'ETUDE</b> .....	<b>23</b>
3.1 <b>Population et territoire</b> .....	<b>23</b>
3.2 <b>Défaveur sociale</b> .....	<b>26</b>
3.3 <b>Offres de soins de premier recours</b> .....	<b>28</b>
3.4 <b>Proxy du tabagisme</b> .....	<b>29</b>
3.5 <b>Corrélation des variables d'ajustement</b> .....	<b>31</b>
<b>4 RÉSULTATS</b> .....	<b>32</b>
4.1 <b>Caractéristiques sociodémographiques et histologiques des cas</b> .....	<b>32</b>
4.1.1 Cancers urothéliaux .....	32
4.1.2 Cancers du rein.....	34
4.1.3 Leucémies aiguës .....	35
4.2 <b>Incidence départementale et répartition spatiale des cas</b> .....	<b>38</b>
4.2.1 Cancers de la vessie .....	38
4.2.2 Cancers du rein.....	46
4.2.3 Leucémies aiguës myéloïdes .....	49

<b>5</b>	<b>DISCUSSION .....</b>	<b>53</b>
<b>5.1</b>	<b>Synthèse des résultats .....</b>	<b>53</b>
<b>5.2</b>	<b>Cancers de la vessie.....</b>	<b>53</b>
5.2.1	Caractéristiques sociodémographiques et histologiques .....	53
5.2.2	Cartographie des risques et détection de clusters : .....	54
<b>5.3</b>	<b>Cancers du rein.....</b>	<b>56</b>
5.3.1	Caractéristiques sociodémographiques et histologiques .....	56
5.3.2	Cartographie des risques et détection de clusters : .....	56
<b>5.4</b>	<b>Leucémies aiguës.....</b>	<b>57</b>
5.4.1	Caractéristiques sociodémographiques et histologiques .....	57
5.4.2	Cartographie des risques et détection de clusters .....	58
<b>5.5</b>	<b>Discussion des choix méthodologiques.....</b>	<b>59</b>
5.5.1	Choix des variables d'ajustement.....	59
5.5.2	Lissage spatial des risques .....	61
5.5.3	Mesure de l'agrégation et détection de clusters.....	63
<b>6</b>	<b>CONCLUSIONS .....</b>	<b>66</b>
	<b>Références bibliographiques .....</b>	<b>67</b>
	<b>Annexe 1 : Regroupements d'Iris peu ou pas peuplés .....</b>	<b>77</b>
	<b>Annexe 2 : Établissements organisés par RCP, par secteur géographique et par statut.....</b>	<b>79</b>
	<b>Annexe 3 : Modèle de lissage spatial BYM.....</b>	<b>81</b>
	<b>Annexe 4 : Méthodes de détection de clusters .....</b>	<b>86</b>
	<b>Annexe 5 : Répartition des cas par sexe et classe d'âge .....</b>	<b>89</b>
	<b>Annexe 6 : Description des tumeurs urothéliales.....</b>	<b>91</b>
	<b>Annexe 7 : Description des tumeurs rénales .....</b>	<b>92</b>
	<b>Annexe 8 : Description des leucémies aiguës .....</b>	<b>93</b>
	<b>Annexe 9 : Description histologique des leucémies aiguës .....</b>	<b>94</b>
	<b>Annexe 10 : Résultats des modèles de lissage.....</b>	<b>95</b>
	<b>Annexe 11 : Analyse de sensibilité des refus.....</b>	<b>97</b>
	<b>Annexe 12 : Résultats de l'analyse de sensibilité des LAM .....</b>	<b>102</b>

## Abréviations

<b>3C</b>	Centre de coordination en cancérologie
<b>ACP</b>	Anatomo-cytopathologie
<b>ACS</b>	Autocorrélation spatiale
<b>AIC</b>	Critère d'information d'Akaike
<b>AP-HM</b>	Assistance publique - Hôpitaux de Marseille
<b>APL</b>	Accessibilité potentielle localisée
<b>ARS</b>	Agence régionale de santé
<b>BYM</b>	(Modèle de) Besag, York et Mollié
<b>CIM10</b>	Classification internationale des maladies 10 <sup>e</sup> révision
<b>CIMO3</b>	Classification internationale des maladies pour l'oncologie 3 <sup>e</sup> révision
<b>CIRC</b>	Centre international pour la recherche sur le cancer
<b>CLCC</b>	Centre de lutte contre le cancer
<b>DIC</b>	<i>Deviance information criterion</i>
<b>Drees</b>	Direction de la recherche, des études, de l'évaluation et des statistiques
<b>EBI</b>	<i>Empirical Bayes Index</i> (de Moran)
<b>EBV</b>	Virus d'Epstein-Barr
<b>EDI</b>	<i>European deprivation index</i>
<b>EGIL</b>	<i>European Group for Immunology of Leukemia</i>
<b>FAB</b>	Franco-americano-britannique
<b>Fdep</b>	<i>French deprivation index</i>
<b>GPS</b>	<i>Global Positioning System</i>
<b>HAP</b>	Hydrocarbures aromatiques polycycliques
<b>HCSF</b>	Haut Conseil de la santé publique
<b>HTLV</b>	<i>Human T-cell Leukemia/Lymphoma Virus</i>
<b>IDS</b>	Indice de défaveur sociale
<b>IGN</b>	Institut national de l'information géographique et forestière
<b>Insee</b>	Institut national de la statistique et des études économiques
<b>Inserm</b>	Institut national de la santé et de la recherche médicale
<b>InVS</b>	Institut de veille sanitaire
<b>IPC</b>	Institut Paoli-Calmettes
<b>Irdes</b>	Institut de recherche et documentation en économie de la santé
<b>Iris</b>	Îlots regroupés pour l'information statistique
<b>IRM</b>	Imagerie par résonance magnétique
<b>Isspam</b>	Institut des sciences de la santé publique de l'Université Aix Marseille
<b>LAM</b>	Leucémie aiguë myéloïde
<b>LAL</b>	Leucémie aiguë lymphoïde (ou lymphoblastique)
<b>MPO</b>	Myélopéroxydase
<b>OMS</b>	Organisation mondiale de la santé
<b>ONU</b>	Organisation des Nations unies
<b>Orep</b>	Observatoire régional d'épidémiologie pour la région PACA
<b>ORS</b>	Observatoire régional de la santé
<b>P2RT</b>	Plan régional de réduction du tabagisme

<b>PACA</b>	Provence-Alpes-Côte d'Azur
<b>PCB</b>	PolyChloroBiphényle
<b>PMSI</b>	Programme de médicalisation des systèmes d'information
<b>RCC</b>	Carcinome à cellules rénales
<b>RCP</b>	Réunion de concertation pluridisciplinaire
<b>RR</b>	Risque relatif
<b>RNCE</b>	Registre national des cancers de l'enfant
<b>RTUV</b>	Résection transurétrale de vessie
<b>Sesstim</b>	Sciences économiques et sociales de la santé et traitement de l'information médicale
<b>SIR</b>	Ratio d'incidence standardisée
<b>TNM</b>	<i>Tumour, Nodes, Metastasis</i>
<b>TVES</b>	Tumeur de la voie excrétrice supérieure
<b>TVIM</b>	Tumeur infiltrant le muscle
<b>TVNIM</b>	Tumeur n'infiltrant pas le muscle
<b>UICC</b>	Union internationale contre le cancer
<b>UMR</b>	Unité mixte de recherche
<b>VIH</b>	Virus de l'immunodéficience humaine

# 1 CONTEXTE ET OBJECTIFS

## 1.1 Contexte local

C'est pour répondre à une saisine de l'Agence régionale de santé (ARS) de la région Provence-Alpes-Côte d'Azur (PACA), que l'Observatoire REVELA13 a été créé en 2012 par la Cellule régionale de Santé publique France en PACA. Cet observatoire des cancers du rein, de la vessie et des leucémies aiguës de l'adulte situé dans le département des Bouches-du-Rhône est un dispositif expérimental de surveillance de cancers ciblés à une échelle territoriale. Il est financé par l'ARS PACA et Santé publique France qui assure également la coordination scientifique.

Pour rappel, le choix des cancers à surveiller a été fait à partir des résultats de différentes études réalisées localement, des spécificités industrielles du département et de rapports nationaux sur le thème cancers et environnements (1,2), en tenant compte des moyens alloués. Entre 2005 et 2011, la Cellule régionale a investigué deux clusters sur le département des Bouches-du-Rhône – un cluster d'hémopathies malignes chez des enfants et un cluster de cancers du rein en milieu professionnel (3) – et a conduit une étude écologique sur le pourtour de l'étang de Berre (à l'ouest de Marseille), montrant un excès d'hospitalisations pour les leucémies aiguës de l'adulte chez l'homme (4). Le ressenti de médecins exerçant sur cette zone quant à l'augmentation croissante des cancers de la vessie a également été pris en compte. Ces trois cancers avaient également été identifiés parmi les dix potentiellement liés à l'environnement selon les critères établis par le département santé-environnement de l'Institut de veille sanitaire – InVS (2).

Depuis 2012, l'Observatoire REVELA13 assure en continu le recueil et l'enregistrement des cas incidents de ces trois cancers sur l'ensemble du département des Bouches-du-Rhône. Le dispositif a pour fonctions :

- 1) d'étudier la répartition de ces trois cancers à différentes échelles géographiques ;
- 2) de suivre l'évolution de cette répartition dans le temps et dans l'espace ;
- 3) de détecter d'éventuels agrégats spatio-temporels de cas.

Ces données pourront par la suite être utilisées pour réaliser des études écologiques environnementales ou des études en milieu professionnel.

Après cinq années de recueil de données et après avoir testé, vérifié et amélioré l'exhaustivité du dispositif, des premières analyses spatiales ont été conduites sur les données d'incidence de 2013 à 2016 (5,6). Les résultats ont montré que l'incidence départementale estimée du cancer de la vessie chez les personnes de 20 ans et plus était significativement supérieure à l'incidence nationale pour les deux sexes. Chez les hommes, une surincidence des cancers de la vessie était identifiée dans la partie sud-est du département avec deux clusters significatifs, le premier incluant six arrondissements de Marseille et le deuxième incluant deux arrondissements de Marseille et 22 communes au sud-est de la ville. Chez les femmes, aucune surincidence significative localisée des cancers de la vessie n'était retrouvée. En revanche, un cluster significatif, incluant neuf arrondissements de Marseille et une commune située à l'est de la ville était identifié. Les différences de répartition géographique observées à l'échelle de la commune n'étaient expliquées par la défaveur sociale ni chez l'homme ni chez la femme. Cependant, cette répartition était liée à l'accès à un urologue pour les femmes.

Les incidences départementales estimées du cancer du rein et des leucémies aiguës myéloïdes (LAM) chez les personnes de 20 ans et plus n'étaient pas différentes des incidences

nationales pour les deux sexes. Les niveaux d'incidence de cancer du rein étaient répartis de façon homogène sur le territoire et aucun cluster n'était identifié. Les mêmes résultats étaient observés pour les LAM. Ces derniers résultats portaient cependant sur un faible nombre de cas.

La convergence des résultats obtenus par différentes méthodes et la cohérence avec ceux issus de la littérature étaient encourageantes. Cependant, ces premiers résultats devaient être consolidés avec des données supplémentaires, notamment en ce qui concernait les cancers du rein et les leucémies aiguës de l'adulte. En effet, ces derniers étant plus rares, la distribution des cas par commune montrait un grand nombre d'entre elles sans cas, limitant ainsi la portée des analyses spatiales. De plus, la détection de regroupement de cas à l'échelle de la commune, pour les cancers plus fréquents de la vessie, ne permettait pas d'attester de l'existence éventuelle de clusters nichés au sein d'une commune et des analyses à une échelle géographique plus fine étaient envisagées pour tester cette hypothèse. Enfin, les surincidences observées pour le cancer de la vessie chez les hommes interrogeaient sur la non-prise en compte d'autres facteurs de confusion, notamment le tabagisme, principal facteur de risque pour ce cancer<sup>1</sup>.

## 1.2 Objectifs

Les objectifs de cette nouvelle étude étaient de conforter les premiers résultats des analyses spatiales conduites en 2019 avec deux années d'incidence supplémentaires (2017 et 2018), de réaliser la démarche de recherche d'agrégats de cancers à une échelle géographique plus fine (Iris) et de produire des représentations cartographiques ajustées sur des cofacteurs pouvant influencer sur la répartition spatiale des cancers (densité de population, accès aux soins, défaveur sociale et tabagisme). Ces cofacteurs peuvent en effet jouer un rôle de facteurs de confusion dans la recherche de facteurs d'exposition environnementale.

---

<sup>1</sup> <https://www.cancer-environnement.fr/225-Cancer-de-la-vessie.ce.aspx>

## 2 MATÉRIEL ET MÉTHODES

Ce travail a été mené en partenariat avec l'unité mixte de recherche en sciences économiques et sociales de la santé et traitement de l'information médicale (UMR1252 Sesstim).

### 2.1 Période et zone d'étude

Cette nouvelle analyse a porté sur les données consolidées de l'Observatoire REVELA13 pour la période 2013-2018.

La zone d'étude concernait le département des Bouches-du-Rhône. Les cas de cancers ciblés ont été agrégés à l'échelle de la commune (arrondissement pour la commune de Marseille) et à l'échelle des îlots regroupés pour l'information statistique (Iris), selon les critères de définition de l'Institut national de la statistique et des études économiques (Insee).

### 2.2 Populations et définitions de cas

Les définitions de cas utilisées pour leur recensement et leur description, basées sur la pratique des médecins urologues, étaient différentes de celles utilisées pour les estimations d'incidence et les analyses spatiales, ces dernières étant celles utilisées par les épidémiologistes et les registres de cancer.

Ainsi, la description clinique et histologique des cancers urothéliaux regroupait l'ensemble des tumeurs de vessie et du haut appareil urinaire alors que celle des cancers du rein portait uniquement sur les tumeurs du parenchyme rénal. En revanche, les calculs d'incidence et les analyses spatiales portaient d'une part sur les cancers de vessie et d'autre part sur les cancers du rein au sens large, incluant les cancers du parenchyme rénal et les tumeurs du haut appareil urinaire, selon la définition utilisée en Europe et au niveau national pour calculer l'incidence de cette localisation.

Pour les leucémies aiguës de l'adulte, l'incidence nationale de référence n'était disponible que pour les LAM, qui représentent la grande majorité des leucémies aiguës incidentes de l'adulte. De ce fait, les estimations d'incidence et les analyses spatiales étaient restreintes aux LAM alors que le recensement et la description des cas concernaient toutes les leucémies aiguës incidentes recensées.

#### 2.2.1 Recensement des cas

La population cible était constituée de toute personne adulte (âgée de 18 ans et plus), résidant dans le département des Bouches-du-Rhône.

Les cas recensés étaient issus de trois sources de données :

- les réunions de concertation pluridisciplinaire (RCP) qui valident la prise en charge et le programme de soins de chaque nouveau patient atteint de cancer ;
- le programme de médicalisation des systèmes d'information (PMSI) des établissements de santé ;
- les comptes rendus d'histologie des laboratoires de cyto-anatomopathologie (ACP) qui examinent les prélèvements de tissus cancéreux.

Dans une étude menée en 2018, la prise en compte de ces trois sources de données complémentaires a permis d'estimer l'exhaustivité globale de l'Observatoire REVELA13 à 99 % pour les tumeurs urothéliales, 98 % pour les cancers du rein et 96 % pour les leucémies aiguës, et de corriger les différences territoriales identifiées (5).

Lors de la mise en place de l'Observatoire en 2012, le recensement des cas reposait uniquement sur le signalement des patients par les médecins urologues et onco-hématologues à l'origine du diagnostic. Les médecins volontaires rejoignaient un réseau de spécialistes en signant une charte d'adhésion les engageant à signaler leurs patients, susceptibles de répondre aux définitions de cas, soit directement à l'Observatoire, soit au travers des RCP. L'ensemble des médecins contactés avait accepté de participer à ces réseaux à l'exception de quatre urologues sur la période d'étude considérée. Parmi ces derniers, deux exercent à Carpentras, un à Marseille et le dernier à Martigues. Les patients pris en charge par ces derniers, ne sont pas inclus dans la base de données de l'Observatoire et étaient donc exclus des analyses. Cependant, des informations concernant le sexe, l'âge au diagnostic, le type de cancer, ainsi que la commune de résidence de ces patients ont été collectées par d'autres sources médicales intervenant dans leur prise en charge. Ces informations ont été utilisées pour réaliser des analyses de sensibilité.

Les critères d'inclusion ont été vérifiés pour chaque cas recensé et une fiche de signalement a ensuite été complétée avec des données démographiques, cliniques et histologiques. Chaque cas validé a été enregistré sur la base de données REVELA13. À partir de l'adresse précise des patients au moment du diagnostic, les cas ont été géolocalisés par l'attribution automatisée de coordonnées géographiques (GPS).

Afin de décrire la répartition des cas en termes d'origine du diagnostic, la RCP et l'établissement à l'origine du diagnostic, ainsi que leurs secteurs géographiques ont été recensés pour chaque cas. Le statut de l'établissement était également spécifié : public, privé ou faisant partie des centres de lutte contre le cancer (Institut Paoli Calmettes de Marseille, Institut Sainte-Catherine d'Avignon et l'Institut du cancer de Montpellier).

Pour réaliser les comparaisons des incidences du département avec celles de la France métropolitaine sur la période d'étude considérée, tous âges confondus, le nombre de cas observés chez les enfants de 0 à 18 ans pour chacun des cancers étudiés par tranche d'âge quinquennal et par sexe ont été obtenus auprès du Registre national des cancers de l'enfant (RNCE).

## 2.2.2 Gestion des données nominatives

Le traitement des données nominatives de l'Observatoire REVELA13 est sous la responsabilité de Santé publique France. L'agence a reçu pour ce traitement, fondé sur l'exercice d'une mission d'intérêt public, une autorisation<sup>2</sup> de la Commission nationale informatique et libertés (Cnil). Ce traitement n'a pas de caractère obligatoire. Conformément à la loi du 6 janvier 1978<sup>3</sup>, les patients disposent d'un droit d'accès, de rectification, de suppression, de limitation et d'opposition à la transmission de leurs données, à leur conservation et à leur analyse.

Les informations recueillies sont couvertes par le secret médical et sont conservées et exploitées dans des conditions assurant leur sécurité et le respect de la vie privée des personnes. Seul le personnel de l'Observatoire REVELA13 habilité par l'investigateur principal a accès, dans la limite nécessaire à la réalisation de leur mission, aux données ainsi collectées. Enfin, les données sont saisies sur une base de données hébergée par Epiconcept®, hébergeur agréé de données de santé.

<sup>2</sup> Autorisation de la Commission nationale informatique et libertés (N° 911545).

<sup>3</sup> Loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Articles 49 et suivants.

## 2.2.3 Définitions de cas

### 2.2.3.1 Définition de cas pour les analyses descriptives

Les définitions de cas retenus pour les descriptions cliniques et histologiques sont :

#### **Cas incident de cancer urothélial**

Un nouveau cas de cancer urothélial correspondait à un premier diagnostic de tumeur invasive de la vessie codée selon la Classification internationale des maladies 10<sup>e</sup> révision (CIM10) (C67), de l'uretère (C66), du bassinet (C65) ou de l'urètre ou autres localisations (C68), diagnostiqué entre le 1<sup>er</sup> janvier 2013 et le 31 décembre 2018. Les tumeurs retenues dans l'étude étaient les tumeurs classées pT1 à pT4 de la classification TNM (*Tumour, Nodes, Metastasis*) de l'Union internationale contre le cancer (UICC).

#### **Cas incident de cancer du rein**

Un nouveau cas de cancer du rein correspondait à un premier diagnostic de tumeur invasive du parenchyme rénal confirmé par histologie (CIM10 : C64) entre le 1<sup>er</sup> janvier 2013 et le 31 décembre 2018.

#### **Cas incident de leucémie aiguë**

Un nouveau cas de leucémie aiguë correspondait à un premier diagnostic de leucémie établi entre le 1<sup>er</sup> janvier 2013 et le 31 décembre 2018 et répondant aux codes issus de la classification CIM10 suivants : C910, C918, C920, C923, C924, C925, C926, C928, C930, C940, C942, C944, C947, C950, C835, C837. Les cas identifiés ont été ensuite recodés selon la Classification internationale des maladies pour l'oncologie 3<sup>e</sup> révision (CIM03), classification retenue pour le calcul de l'incidence nationale.

### 2.2.3.2 Définition de cas pour les calculs d'incidences

Les définitions de cas utilisées pour les calculs d'incidence sont basées sur celles utilisées par les registres de cancer :

#### **Cas incident de cancer de vessie**

Un cas incident de cancer de vessie était un cas de cancer urothélial dont le code CIM10 est égal à C67.

#### **Cas incident de cancer du rein**

Les cas incidents de cancer du rein regroupaient les adénocarcinomes du parenchyme rénal (C64) et les tumeurs malignes des voies excrétrices, à l'exception de la vessie, (C65-C66 et C68).

#### **Cas incident de LAM**

Un cas incident de LAM correspondait à un diagnostic de leucémie aiguë dont le code CIM03 correspond à l'un des codes suivants : 97273, 98403, 98613, 98653, 98663, 98673, 98693, 98703, 98713, 98723, 98733, 98743, 98713, 98773, 98783, 98793, 98913, 98953, 98963, 98973, 98981, 98983, 99103, 99113, 99123, 99203, 99303, 99313.

## 2.2.4 Données de population

### 2.2.4.1 Sources des données

Les données de population annuelle à l'échelle des départements, des communes et des Iris étaient issues des bases de données de l'Insee de 2013 à 2018 inclus. Les données des populations départementales et communales étaient disponibles par sexe et âge révolu.

Les données des populations des Iris désagrégées par sexe n'étaient disponibles que par classes d'âge de 15 ans.

#### 2.2.4.2 Calcul du nombre de personnes-années

Les taux d'incidence ont été calculés par classe d'âge en effectuant le ratio du nombre total de cas incidents sur le nombre total de personnes-années sur la période d'étude.

À l'échelle de la commune (arrondissement pour Marseille), les personnes-années 2013-2018 ont été calculées en âge révolu, à partir des données de population de l'Insee regroupées par commune, sexe et âge annuel.

À l'échelle des Iris, les données de populations par sexe étant fournies par classes d'âge de 15 ans, il n'a pas été possible de calculer précisément un nombre de personnes-années. Ces effectifs ont été approximés par la population pour chacune des classes d'âge.

Par ailleurs, les analyses spatiales s'effectuant sur les cas âgés de 20 ans et plus, la population des Iris de la classe d'âge [20 ; 29] a été estimée en utilisant le ratio de population de la classe d'âge [20 ; 29] dans la classe d'âge [15 ; 29] dans la commune de chaque Iris.

Les données de population, à l'échelle communale comme à l'Iris, étaient disponibles pour les années 2013 à 2018 au moment de l'analyse, mais pas celles de l'année 2019. Les personnes années 2019 ont été estimées en supposant que les populations 2019 étaient identiques aux populations 2018.

## 2.3 Données géographiques

### 2.3.1 Sources des données

Le fichier de données géographique des communes était issu des données Admin Express de l'Institut national de l'information géographique et forestière (IGN) dans sa mise à jour de décembre 2021<sup>4</sup>.

Le fichier de données géographique des Iris du département des Bouches-du-Rhône était issu des données Iris\_GE de l'Insee et l'IGN dans sa mise à jour du 1<sup>er</sup> janvier 2021<sup>5</sup>.

### 2.3.2 Regroupement des Iris peu ou pas peuplés

La réalisation des analyses spatiales nécessitait de disposer d'une population suffisante dans chaque unité géographique (ici commune ou Iris). En effet, le nombre de cas attendus dans chaque unité était calculé à partir de sa population résidente.

Les Iris non peuplés (Iris d'activité ou « divers ») ont créé des « trous » engendrant une discontinuité géographique, ce qui a posé un problème lors des analyses car la méthode de lissage spatial utilisée est basée sur la contiguïté des unités spatiales. Les Iris peu peuplés, définis par un seuil de 25 hommes ou 25 femmes, ont été fusionnés avec l'Iris le plus proche, de centre à centre. Les populations des Iris regroupés ont ensuite été sommées et les variables d'ajustements utilisées étaient celles de l'Iris regroupé.

---

<sup>4</sup> <https://geoservices.ign.fr/adminexpress>

<sup>5</sup> <https://geoservices.ign.fr/Irisge>

Au total, 19 Iris peu peuplés, dont 13 dans la commune de Marseille, ont été regroupés avec les 17 Iris les plus proches. De plus, l'Iris « Les îles » situé au large de Marseille a été regroupé avec l'Iris de « Malmousque-Pêcheurs » afin d'assurer une continuité territoriale.

La carte et la liste des regroupements sont présentées en Annexe 1.

### 2.3.3 Identification des communes et Iris de résidence des cas

La commune de résidence des cas a été déterminée par l'adresse postale de résidence enregistrée au moment du diagnostic. Les coordonnées géographiques correspondant à l'adresse postale ont été utilisées pour déterminer l'Iris de résidence par l'identification des frontières d'Iris dans lequel chaque point GPS se situait. Pour deux cas, le point GPS n'était pas renseigné, rendant impossible l'attribution d'un Iris. Ces cas ont été exclus de l'analyse par Iris.

Pour 21 cas, l'Iris de résidence ainsi déterminé ne correspondait pas à leur commune de résidence renseignée par l'adresse postale. La majeure partie du temps, ces divergences étaient dues au passage des frontières d'Iris au milieu des rues, et à l'imprécision des points GPS. Les distances les plus courtes des points GPS aux frontières d'un Iris des communes de résidence variaient dans ces cas de 0 à 22 mètres. L'Iris correspondant a été attribué. Pour un cas, la commune de résidence se situait au-delà de 3 000 mètres. Ce cas a été exclu de l'analyse par Iris.

## 2.4 Facteurs d'ajustement

Les facteurs pris en compte dans cette étude recouvraient à la fois des facteurs de confusion, tels que les inégalités sociales, la densité de population ou l'accès aux soins, et des facteurs de risque tel que le tabagisme.

### 2.4.1 Densité de population

Bien que la densité de population n'influence pas directement l'incidence des cancers comme elle le fait pour des maladies transmissibles, elle est néanmoins corrélée à de multiples facteurs de risques de cancers (variables sociodémographiques, mode de vie, concentration de polluants, type d'habitat, etc.). Elle a été considérée dans notre étude comme une variable d'ajustement.

La densité de population reflète l'occupation moyenne du territoire sans tenir compte d'une éventuelle hétérogénéité dans la répartition des habitants avec par exemple un centre-ville densément peuplé et un territoire quasi vide autour. Pour limiter cet effet négatif, la densité de population a été remplacée par la grille de densité communale définie par l'Insee. Basée sur une méthodologie européenne, la caractérisation des communes a été faite sur la proportion d'habitants résidant dans des zones densément peuplées (mailles de 1 km<sup>2</sup>) au sein de la commune (7). Quatre niveaux ont été ainsi définis : communes densément peuplées, communes de densité intermédiaire, communes rurales peu denses et communes rurales très peu denses. Les données disponibles au 1<sup>er</sup> janvier 2022 ont été utilisées.

L'échelle de l'Iris n'est pas soumise à la même problématique puisque par définition les Iris sont « homogènes quant au type d'habitat » (8). De plus, leur taille restreinte limite généralement la variabilité de l'habitat. Pour les analyses à l'échelle de l'Iris, la densité de population a donc été calculée à partir de la population moyenne en personnes-années et de

l'aire des Iris ou Iris regroupés, puis discrétisée en quartiles pour assurer des effectifs suffisants dans chaque classe ainsi créée.

## 2.4.2 Indices de défaveur sociale

En 2019, l'Organisation des Nations unies (ONU) a rappelé que les inégalités sociales avaient un impact important à tous les stades du cancer et de la vie<sup>6</sup>. En effet, les différences entre les groupes sociaux affectent non seulement l'exposition d'un individu aux facteurs de risque et à la probabilité de développer un cancer, mais aussi l'accès aux installations de dépistage, de diagnostic et de traitement, y compris aux soins palliatifs. En France, une étude menée sur plusieurs localisations cancéreuses a montré qu'il existait un gradient d'incidence croissant statistiquement significatif avec la défaveur sociale pour de nombreuses localisations et notamment pour le cancer de la vessie chez l'homme (9). La défaveur sociale a donc été prise en compte dans notre étude comme un facteur de confusion dans la recherche des facteurs d'exposition au travers des analyses spatiales.

Les variables sociales individuelles pour chaque cas n'étant pas disponibles dans la base de données REVELA13, une approximation de la position socio-économique des cas a été réalisée en affectant un indice de défaveur sociale (IDS) correspondant à sa commune ou à son Iris de résidence au moment du diagnostic.

En 2019, lors de la réalisation des premières analyses spatiales, trois indicateurs de défaveur sociale ont été utilisés : l'*European deprivation index* (EDI), l'indice de Townsend et le *French deprivation index* (Fdep), sans que les études de sensibilité ne mettent en évidence de différence significative dans les résultats obtenus. Compte tenu du mode de construction du Fdep, à partir de données agrégées à la commune, et de l'adéquation de sa distribution à cette échelle avec la connaissance empirique de la défaveur sociale sur le département, le Fdep a été considéré comme l'IDS le mieux adapté à cette nouvelle étude.

Le Fdep est construit à partir des données du recensement et des revenus fiscaux :

- le revenu fiscal médian par unité de consommation ;
- la part des diplômés de niveau baccalauréat (minimum) dans la population de 15 ans ou plus non scolarisée ;
- la part des ouvriers dans la population active de 15 à 64 ans ;
- la part des chômeurs dans la population active de 15 à 64 ans (10,11).

Les données de la version mise à jour en 2015 du Fdep ont été utilisées pour cette étude (12).

## 2.4.3 Accès aux soins de premier recours

L'offre de soins sur le département des Bouches-du-Rhône n'est pas homogène du fait de la disparité des profils de territoire ; elle peut donc influencer sur la détection précoce et la prise en charge des cancers. L'accès à l'offre de soins a donc également été pris en compte comme facteur de confusion potentiel dans les analyses spatiales.

En 2019, un indicateur d'accès à l'offre de soins de spécialistes en urologie avait été construit faute d'indicateurs validés disponibles en France. Cet indicateur est toutefois difficile à mettre à jour et peu discriminant car la plupart des urologues exercent à Marseille. Pour cette nouvelle étude, le choix s'est porté sur un indicateur largement utilisé par ailleurs et régulièrement mis à jour : l'accessibilité potentielle localisée (APL). L'APL a été développé par la Direction de la recherche, des études, de l'évaluation et des statistiques (Drees) et l'Institut de recherche et

---

<sup>6</sup> <https://unric.org/fr/sante-lonu-appelle-a-reduire-les-inegalites-sociales-face-au-cancer/>

documentation en économie de la santé (Irdes) pour mesurer l'adéquation spatiale entre l'offre et la demande de soins de premier recours (médecine générale) à un échelon géographique fin (13).

Cet indicateur n'étant pas disponible à l'échelle de l'Iris, la valeur de l'APL de sa commune d'appartenance a été attribuée à chaque Iris. Les données utilisées correspondent à la mise à jour de 2018 (13).

## 2.4.4 Tabagisme

Le tabagisme est un facteur de risque de nombreux cancers et notamment le principal facteur de risque du cancer de la vessie. En 2017, la prévalence standardisée du tabagisme quotidien en région PACA s'élevait à 32,2 % parmi les 18-75 ans alors que la moyenne nationale était de 26,9 % en France métropolitaine (14). En l'absence de données de prévalence du tabagisme à l'échelle de la commune, une rapide revue de la littérature a montré que d'autres approches étaient utilisées pour construire un *proxy* du tabagisme. Ainsi, plusieurs auteurs ont utilisé le taux communal normalisé de mortalité par cancer du poumon comme *proxy* du tabagisme pour ajuster leurs analyses (15-17).

Les données de mortalité par cancer du poumon pour les 20 ans et plus ont été obtenues via une extraction du Système national des données de santé (SNDS) pour les communes du département des Bouches-du-Rhône sur la période 2002-2012. Cette période a été choisie pour disposer de données robustes, sur une période antérieure à la période d'étude.

Des ratios standardisés de mortalité (SMR) ont été calculés pour chaque commune à partir du nombre de décès observé rapporté au nombre de décès attendus, obtenus à partir des taux de mortalité par cancer des poumons estimés en France par sexe et par tranche d'âge quinquennal sur la même période de dix ans. Les SMR ont ensuite été lissés au moyen du modèle hiérarchique bayésien de Besag, York et Mollié (BYM) (15).

**Dans les analyses à l'échelle de l'IRIS, sans une meilleure granularité disponible, les indicateurs à l'échelle communale ont été utilisés.**

## 2.5 Analyse descriptive et estimation d'incidence

La description des facteurs d'ajustement présentée sous forme cartographique, a utilisé une discrétisation en quintiles à l'exception de la densité de population, qui est présentée en quatre niveaux comme explicité précédemment. Le découpage en quantiles assure un effectif identique dans chaque classe et a l'avantage de s'appliquer à toutes les distributions selon Rican (18). L'information est donc portée par les bornes définissant ces classes et leur répartition sur le territoire. Les distributions de l'APL, du Fdep et du *proxy* du tabagisme sont globalement continues et assimilables à des distributions normales, aussi la perte d'information sur leur forme due à cette méthode de découpage demeure minimale (18).

Les caractéristiques sociodémographiques, cliniques et histologiques des cas incidents de cancers chez l'adulte de plus de 18 ans, selon les définitions de cas présentées précédemment, ont été décrites chez les hommes et les femmes.

La description histologique des cancers urothéliaux a été réalisée selon le stade et le grade tumoral défini par l'Organisation mondiale de la santé (OMS) en 2016. Le grade correspond au degré de différenciation de la tumeur. Il distingue les tumeurs à faible potentiel de malignité, les tumeurs de bas grade et les tumeurs de haut grade. Les cancers du rein ont été également décrits selon le stade et le grade selon les recommandations de l'OMS, les bas grades

correspondants aux cancers de grade 1 ou 2 et les hauts grades regroupant les grades 3 et 4. Enfin, la répartition des profils de LAM a été réalisée selon la classification de l'OMS (OMS 2008 pour les cas diagnostiqués avant 2016 et OMS 2016 pour les cas diagnostiqués à partir de 2016) et selon la classification Franco-américano-britannique (FAB) lorsque le myélogramme des cas était disponible.

L'analyse descriptive a été complétée par la répartition des cas selon le secteur géographique et le statut (privé, public ou centres de lutte contre le cancer – CLCC) de l'établissement à l'origine du diagnostic. Les établissements de diagnostics des cas étudiés, les RCP associées, leurs secteurs géographiques et leur statut sont présentés en Annexe 2.

Des comparaisons entre hommes et femmes ont été réalisées avec un test du  $\chi^2$  pour les variables catégorielles et avec un test de Student pour les variables continues. Dans le cas où les conditions d'application n'étaient pas respectées, le test Exact de Fisher ou un test non paramétrique de Wilcoxon, ont été utilisés (19).

Des taux d'incidence bruts départementaux pour 100 000 personnes-années ont été calculés pour les hommes et les femmes sur la période cumulée 2013-2018. Les cas observés pour chaque tranche d'âge quinquennal et par sexe ont été obtenus à partir de la somme du nombre de cas incidents âgés de 18 ans et plus enregistrés par l'Observatoire et du nombre de cas observés chez les enfants de 0 à 18 ans, recensé par le Registre national des cancers de l'enfant (RNCE) et mis à disposition de l'Observatoire. Le total des cas observés par tranche d'âge quinquennal et par sexe a ensuite été rapporté aux nombres de personnes-années correspondant sur la même période.

De la même manière, les taux d'incidence cumulée sur la période d'étude pour 100 000 personnes-années par classe d'âge quinquennal et par sexe ont été calculés pour la France métropolitaine à partir des estimations nationales d'incidence de cancer de 2013 à 2018 de Santé publique France (20,21).

Les taux d'incidence standardisés monde par sexe pour 100 000 personnes-années sur la période d'étude ont enfin été calculés en standardisant les taux d'incidence bruts sur l'âge quinquennal, en prenant comme référence la structure de population mondiale publiée en 1976 par le CIRC (22). Les taux d'incidence standardisés monde pour la France ont été également calculés pour comparaison.

## 2.6 Cartographie des risques lissés

La représentation cartographique des indicateurs de cancers permet de décrire leur distribution spatiale. Les indicateurs utilisés pour réaliser ces cartographies sont les ratios d'incidence standardisés (SIR) sur l'âge. Pour chaque unité spatiale (commune ou Iris), le nombre observé de cas incidents pour les patients âgés de 20 ans et plus est comparé au nombre attendu, sous l'hypothèse de taux incidents standardisés par âge et par sexe égaux à ceux de la France métropolitaine (méthode de standardisation indirecte)<sup>7</sup>. Les taux d'incidence cumulée pour la France métropolitaine dérivés des estimations de Santé publique France sont utilisés pour la standardisation.

Le recours à une méthode de lissage permet de diminuer les différences de précision des estimations initiales en partageant l'information apportée par l'ensemble des unités géographiques et ainsi diminuer les excès de risques apparents dus aux faibles effectifs de cas ou de population (23).

---

<sup>7</sup> Les données départementales mises à disposition par le RNCE n'étant pas désagrégées à la commune, elles ne permettaient donc pas d'inclure les classes d'âge quinquennales inférieures à 20 ans dans ces analyses.

Dans la présente mise à jour des analyses, la méthode de lissage spatiale utilisée pour les analyses publiées en 2019 a été utilisée. Il s'agit du lissage bayésien hiérarchique BYM (24). Ce modèle partage le risque résiduel en un effet aléatoire non spatial et un effet spatialement structuré qui suit un modèle gaussien autorégressif conditionnel (25,26). Une présentation plus détaillée de la méthode est présentée en Annexe 3. En suivant les recommandations de Colonna et Sauleau (27,28), l'Annexe 3 présente également le processus de décision qui a mené à la confirmation de l'utilisation du modèle BYM et au choix des paramètres utilisés. La prise en compte de la spatialisation des données nécessitant de définir une structure de voisinage entre les unités spatiales, la matrice de type « Queen » a été appliquée. Deux unités spatiales sont considérées comme voisines si elles ont une frontière ou un sommet en commun.

Les modèles BYM ont été ajustés sur les cofacteurs, introduits en classes pour les variables liées à la densité de population, et en variables continues, pour l'indice de défaveur sociale, l'APL et le proxy du tabagisme.

Une analyse des corrélations entre les facteurs d'ajustement a été réalisée à l'aide des coefficients de corrélation de Pearson. Le cas échéant, une analyse de sensibilité intégrant une interaction entre des facteurs corrélés a été effectuée.

La sélection du modèle s'est basée sur l'analyse de l'autocorrélation des résidus et la minimalisation du *Deviance information criterion* (DIC) (29). Comme pour la première analyse réalisée en 2019, si les variations des DIC différaient de moins de 10 du plus faible DIC, le modèle complet était retenu afin de pouvoir discuter et comparer les effets des différents facteurs d'ajustement (29). L'Annexe 3 détaille les critères utilisés et le processus de choix.

Dans le contexte bayésien, le modèle produit un large ensemble de valeurs pour chaque paramètre permettant d'estimer leurs distributions *a posteriori*. L'estimation ponctuelle des paramètres du modèle correspond à la médiane de ces distributions. Un intervalle de crédibilité à 95 % est calculé en identifiant les percentiles liés à 2,5 % et 97,5 % de ces distributions. Les valeurs des SIR et des effets des différents cofacteurs présentés dans le rapport correspondent à l'exponentielle de chaque médiane encadrée par son intervalle de crédibilité *a posteriori* (voir l'Annexe 3 pour plus de détails).

**Dans l'ensemble du rapport, le terme « SIR lissé » fait référence à l'exponentielle de la valeur médiane et de l'intervalle de crédibilité des distributions *a posteriori* des estimations obtenues après lissage par le modèle BYM. Il en est de même pour les effets des cofacteurs introduits dans les modèles interprétés comme des rapports de risque (RR). La convention d'écriture utilisée dans ce rapport pour l'ensemble des médianes *a posteriori* des effets est : « effet IC95 % [borne inférieure ; borne supérieure] ».**

À partir de cet intervalle de crédibilité, la mesure du risque est estimée selon trois catégories :

- « Surincidence » lorsque la borne inférieure de l'intervalle de crédibilité du SIR lissé est supérieure à 1 ;
- « Sous-incidence » lorsque la borne supérieure de l'intervalle de crédibilité du SIR lissé est inférieure à 1 ;
- « Incidence attendue » quand l'intervalle de crédibilité du SIR lissé inclut 1.

Enfin, le calcul de l'indice de Moran appliqué aux résidus du modèle de lissage retenu permet de vérifier l'absence d'autocorrélation de ces résidus (30).

Des analyses de sensibilité ont été menées :

- En intégrant, uniquement à l'échelle de la commune, le profil des patients (classe d'âge et sexe) pour lesquels les urologues ont refusé qu'ils soient inclus dans la base de données de l'Observatoire (vessie et rein uniquement) ;
- En fusionnant les données des deux sexes afin d'augmenter l'effectif du nombre de cas (LAM uniquement).

Les estimations ponctuelles des SIR lissés sont présentées sous forme de cartes pour visualiser aisément leur répartition sur le territoire. Pour faciliter la lecture de ces cartes, une échelle discrétisée a été utilisée selon l'étendue des estimations pour les trois cancers et en accord avec les seuils utilisés dans les profils régionaux des estimations de rapports de taux d'incidence de cancers publiés par Santé publique France (31). Les valeurs supérieures à 1,6 ont été regroupées dans une même modalité. L'échelle utilise un gradient d'intensité vert-jaune-rouge centré sur la valeur 1 pour visualiser facilement la distribution des estimations. De plus, la surincidence a été représentée par un point noir au centre des unités spatiales concernées sur les cartes présentées.

Pour une localisation de cancer et un sexe donné, l'analyse des risques lissés à l'échelle des Iris a été réalisée uniquement pour un nombre suffisant de cas répartis sur le territoire. Le seuil de 80 % des Iris présentant au moins 1 cas a été utilisé.

Le traitement et l'analyse des données ont été réalisés avec le logiciel R 4.2.0 (32). Pour effectuer la modélisation des modèles BYM, le package *CARBayes* est utilisé (33).

## 2.7 Recherche de clusters

Les méthodes décrites ont été développées pour tester une tendance à l'agrégation des cas d'une pathologie. Elles ont pour objectif de mieux comprendre la distribution géographique des maladies et d'en étudier l'hétérogénéité spatiale (23). Dans cette étude, **un cluster a été défini comme une concentration de cas de cancers anormalement élevée, c'est-à-dire supérieure à celle attendue**, dans une zone géographique du territoire des Bouches-du-Rhône.

Deux types de méthode de détection ont été appliqués dans cette étude : des méthodes globales et une méthode locale de détection de clusters.

### 2.7.1 Méthodes globales de clustering

Ces méthodes évaluent la tendance globale à l'agrégation des cas et à l'hétérogénéité spatiale de leur distribution sans localiser les clusters sur la zone d'étude.

Les quatre indices suivants, parmi les plus utilisés, ont été retenus et appliqués :

- Le test de Potthoff et Whittinghill (PW) qui permet de vérifier l'existence d'une hétérogénéité globale en termes d'instabilité spatiale des SIR (34).
- L'indice de Moran (I), qui résume le degré de ressemblance des unités spatiales voisines et s'interprète comme un coefficient de corrélation (30).
- L'EBI de Moran (*Empirical Bayes Index*) qui est une modification de l'indice de Moran prenant en compte la variabilité de la population des unités spatiales, et qui s'interprète de la même façon (35)
- La statistique de Tango (T) qui teste la tendance au regroupement des cas dans des clusters à l'intérieur de la zone d'étude (36).

Chacun de ces tests renseigne sur une tendance globale à l'agrégation dans la zone d'étude. Néanmoins, même en l'absence de détection d'agrégation globale, des clusters locaux peuvent exister. L'Annexe 4 présente ces méthodes globales en détail.

Ces tests sont disponibles dans les packages R *DCluster* et *spdep*. Les p-valeurs associées à l'indice de Moran et à l'EBI sont obtenues par une procédure de Monte Carlo.

### 2.7.2 Méthode de détection locale non ciblée

L'objectif des méthodes de balayage spatial est la surveillance géographique d'un territoire dans le but de détecter les zones pour lesquelles une incidence plus élevée de cas d'une maladie est observée, sans hypothèse *a priori* (23). Parmi les différentes méthodes développées, la statistique de scan spatial développée par Kulldorff et Nagarwalla est la méthode la plus utilisée et reste incontournable en épidémiologie spatiale (37-39).

Une fenêtre, de forme circulaire ou elliptique centrée sur une unité spatiale, balaye la zone d'étude. Pour chaque taille de fenêtre correspondant à un seuil de population contenue à l'intérieur de cette fenêtre, un test évalue s'il existe un risque plus élevé à l'intérieur de la fenêtre qu'à l'extérieur. Une fois la statistique estimée pour l'ensemble des tailles de clusters potentiels, la fenêtre se déplace sur une autre unité spatiale. Finalement, la fenêtre qui correspond au maximum de vraisemblance est le cluster le plus probable (*i.e.* cluster principal) et une p-valeur, calculée à partir de simulations de Monte Carlo, est assignée à ce cluster de même que le rapport de risque correspondant. D'autres clusters dits clusters secondaires peuvent aussi être identifiés par cette méthode en sélectionnant les valeurs maximales de vraisemblance suivantes des clusters qui ne se superposent pas aux clusters déjà retenus.

Une fenêtre elliptique avec pénalisation, ainsi que l'optimisation de la taille maximale des clusters identifiés par l'utilisation du coefficient de Gini ont été utilisées dans cette étude pour améliorer la détection des clusters de forme irrégulière et la qualité de l'information des clusters identifiés (40-45). Le coefficient de Gini (46) est un résumé de l'information du comportement d'une courbe de Lorenz (47) qui décrit et mesure l'hétérogénéité d'une distribution, originellement utilisée pour décrire la répartition des richesses. Han *et al.* ont adapté ce coefficient à la sélection de clusters détectés par statistique de scan (44). La sélection par le coefficient de Gini (44,45) se base sur le constat que le choix de la taille maximale de la fenêtre de scan influe sur le résultat : une valeur trop grande peut engendrer la détection d'un très grand cluster principal pouvant cacher des clusters moyens plus informatifs ; à l'inverse, une taille trop petite peut entraîner l'éclatement d'un cluster en plusieurs petits clusters moins informatifs. Cette méthode permet de déterminer la taille maximale des clusters identifiés qui maximise l'information contenue dans le jeu de clusters retenu.

La détection de clusters a été réalisée avec une taille maximale de la fenêtre de scan de 30 % de la population pour l'échelle communale et de 10 % de la population pour l'échelle des Iris (42).

Une analyse de sensibilité a été réalisée en excluant la commune de Marseille, qui représente 43 % de la population du département, ce qui pouvait masquer la présence de clusters dans d'autres zones. Pour l'ensemble des tests statistiques, le seuil de significativité statistique a été fixé à 5 %.

La cartographie des clusters identifiés par la méthode de scan spatial présente le cluster principal et les clusters secondaires, le risque relatif (RR) (correspondant au risque dans le cluster par rapport au risque dans le reste de la zone d'étude) et le seuil de significativité.

Les analyses ont été réalisées avec le logiciel SaTScan V10.0.1 (48), en appelant l'exécution du logiciel via R et en utilisant le package *rsatscan*. Les représentations cartographiques ont été réalisées en utilisant les packages *ggplot2* et *sf* du logiciel R.

Une description plus détaillée de l'ensemble des méthodes est disponible en Annexe 4.

# 3 DESCRIPTION DE LA ZONE D'ÉTUDE

## 3.1 Population et territoire

Le département des Bouches-du-Rhône, situé dans la région PACA, est limitrophe des départements du Gard, du Vaucluse et du Var. Il est délimité par le Rhône à l'ouest, la Durance au nord et la mer Méditerranée au sud. Ce territoire, d'une surface de 5 087 km<sup>2</sup>, est largement occupé par des espaces naturels mais est aussi fortement urbanisé et accueille la plus grande métropole française. Ce département comprend 119 communes dont la commune de Marseille, elle-même scindée en 16 arrondissements, soit 134 unités spatiales considérées comme des communes pour notre étude (Figure 1). Il compte 840 Iris dont 393 (47 %) se situent dans les limites de Marseille. Ce nombre a été ramené à 820 Iris après regroupement (voir paragraphe 2.3.2).

Figure 1. Communes du département des Bouches-du-Rhône, 2013-2018



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTM - juin 2022

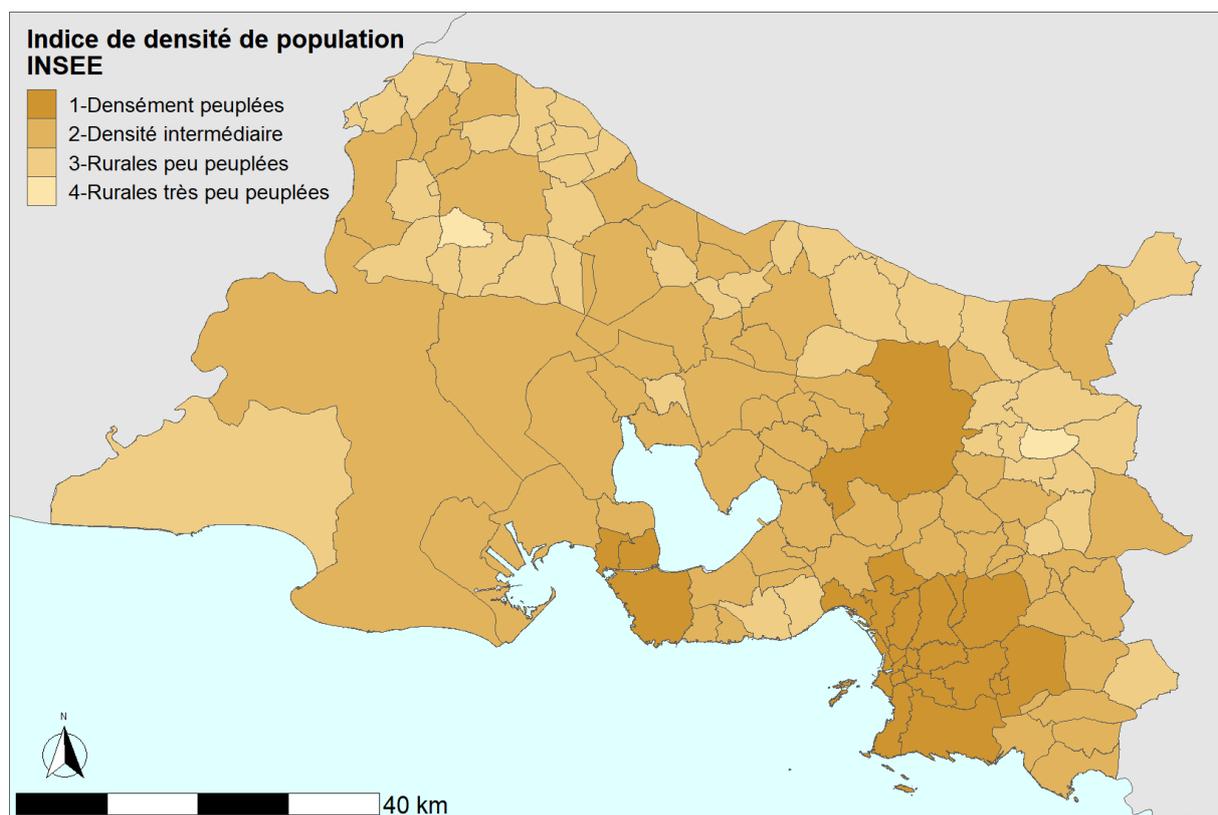
En 2018, le département des Bouches-du-Rhône comptait 2 034 254 habitants, ce qui en faisait le département le plus peuplé de la région (40 % de la population régionale) et le 3<sup>e</sup> le plus peuplé de France. Les communes de Marseille et d'Aix-en-Provence, avec respectivement 868 268 et 143 099 habitants en 2018, concentrent près de la moitié de la

population du département. Le département est également le plus densément peuplé de la région, avec une densité de 394,3 habitants au km<sup>2</sup>.

La population est cependant inégalement répartie sur le territoire : la Camargue et les Alpilles, de même que certaines communes frontalières du Var constituent des zones moins peuplées que le reste du département (Figure 2). Les communes densément peuplées (niveau 1) représentent 17,9 % des communes et 57,6 % de la population alors que les communes rurales peu ou très peu peuplées (niveau 3 ou 4) représentent 34,3 % des communes et 6,2 % de la population.

Une estimation des mouvements de population réalisée par la Direction des statistiques de l'ARS PACA, à partir des données du Système national des données de santé (SNDS) sur la période 2015-2018, montrait que la population des Bouches-du-Rhône était globalement stable sur cette période. En effet, 81 % de la population était restée dans sa commune de résidence, 7 % avaient changé de commune dans le département et 2 % y étaient décédés. Seulement 5 % avaient changé de département et l'information sur les déplacements n'était pas connue pour 5 % de la population (données non publiées).

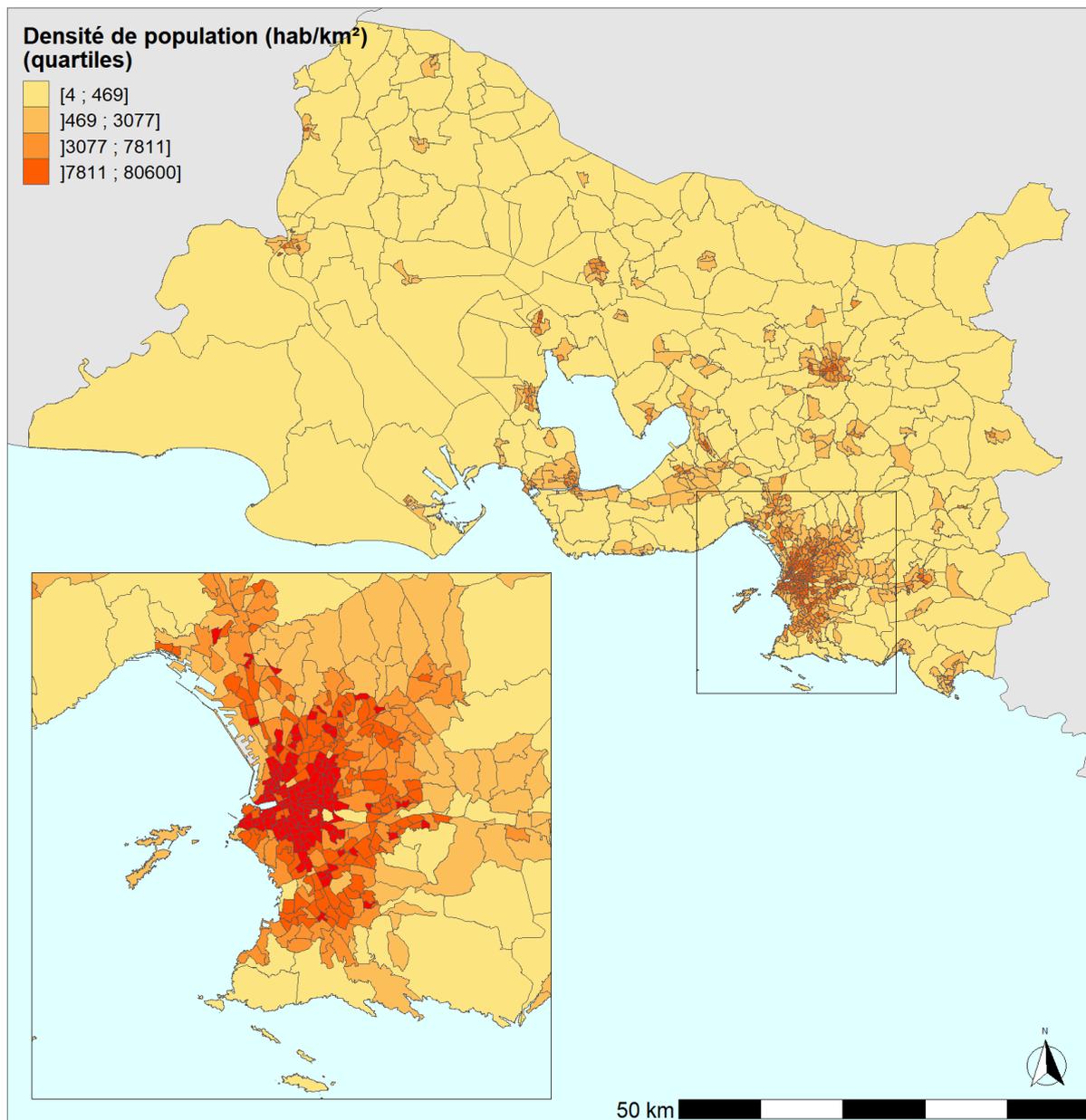
**Figure 2. Indice de densité de l'Insee selon la commune de résidence, Bouches-du-Rhône (données 2022)**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juin 2022

L'échelle de l'Iris, de par sa définition même, permet d'apprécier cette répartition très inégale de la population sur le territoire du département, avec une concentration de population dans des centres-villes ou villages (Figure 3).

**Figure 3. Répartition (quartiles) de la densité de population selon l'IRIS de résidence, Bouches-du-Rhône, moyenne de 2013 à 2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juin 2022

## 3.2 Défavor social

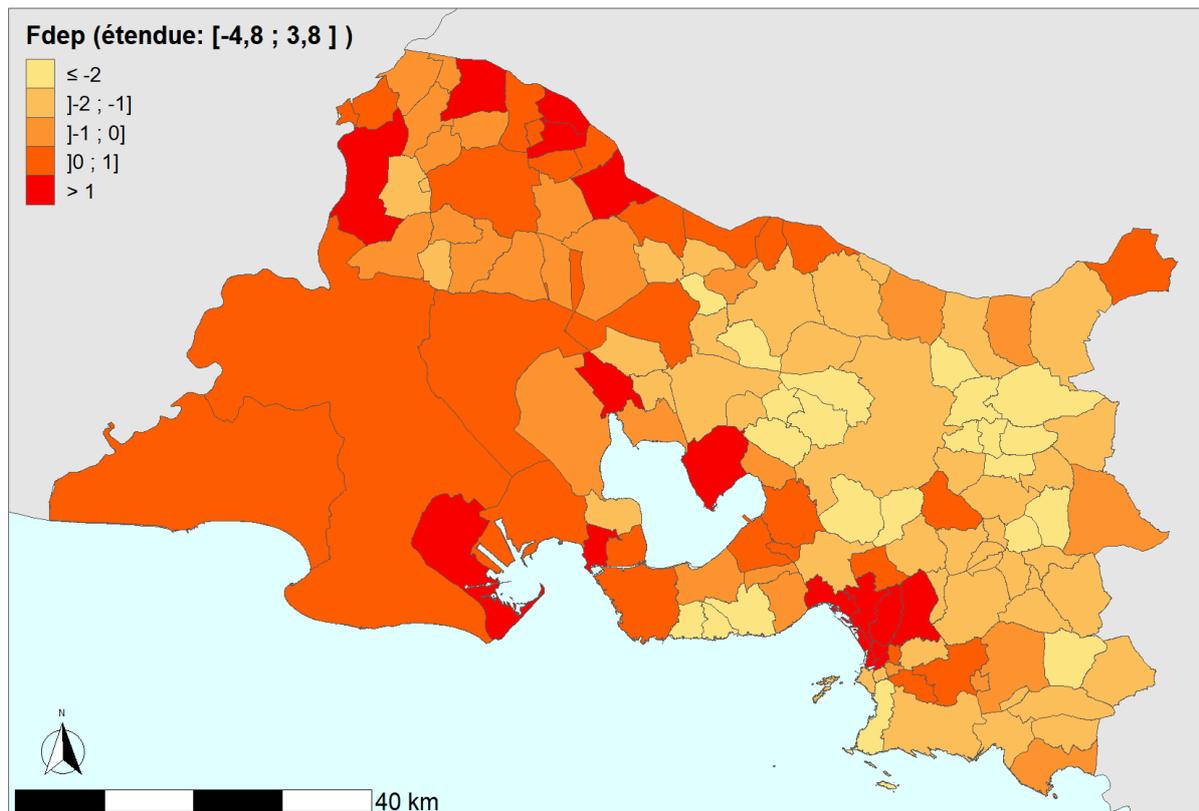
Le département des Bouches-du-Rhône est marqué par la pauvreté et les inégalités sociales. En 2017, 18,2 % des personnes vivaient sous le seuil de pauvreté soit 3,8 points de plus qu'en France métropolitaine. La pauvreté se concentrait dans les quartiers de la politique de ville où résident 16 % des habitants mais 60 % des personnes pauvres du département. Dans certains quartiers de ces communes, plus de la moitié des habitants vivent sous le seuil de pauvreté. Par ailleurs, le rapport interdécile des niveaux de vie (entre le plancher des 10 % les plus riches et le plafond des 10 % les plus pauvres) était de 3,8 (contre 3,4 dans l'Hexagone)<sup>8</sup>.

La distribution de l'indice Fdep à l'échelle de la commune et de l'Iris reflète ces différences, avec un dégradé de couleur, du plus clair pour les unités géographiques dites favorisées, avec un indice faible, au plus foncé pour les unités dites défavorisées, ayant un indice élevé (Figure 4).

Les seuils fixés sont communs aux deux échelles, commune/arrondissement et Iris, pour assurer la comparabilité et se rapprocher du découpage en quintiles pondérés par la population. Les communes à l'est du département sont globalement plus favorisées, à l'exception des arrondissements au nord de Marseille, que celles situées dans la partie ouest.

Figure 4. Répartition de la défaveur sociale représentée par le Fdep

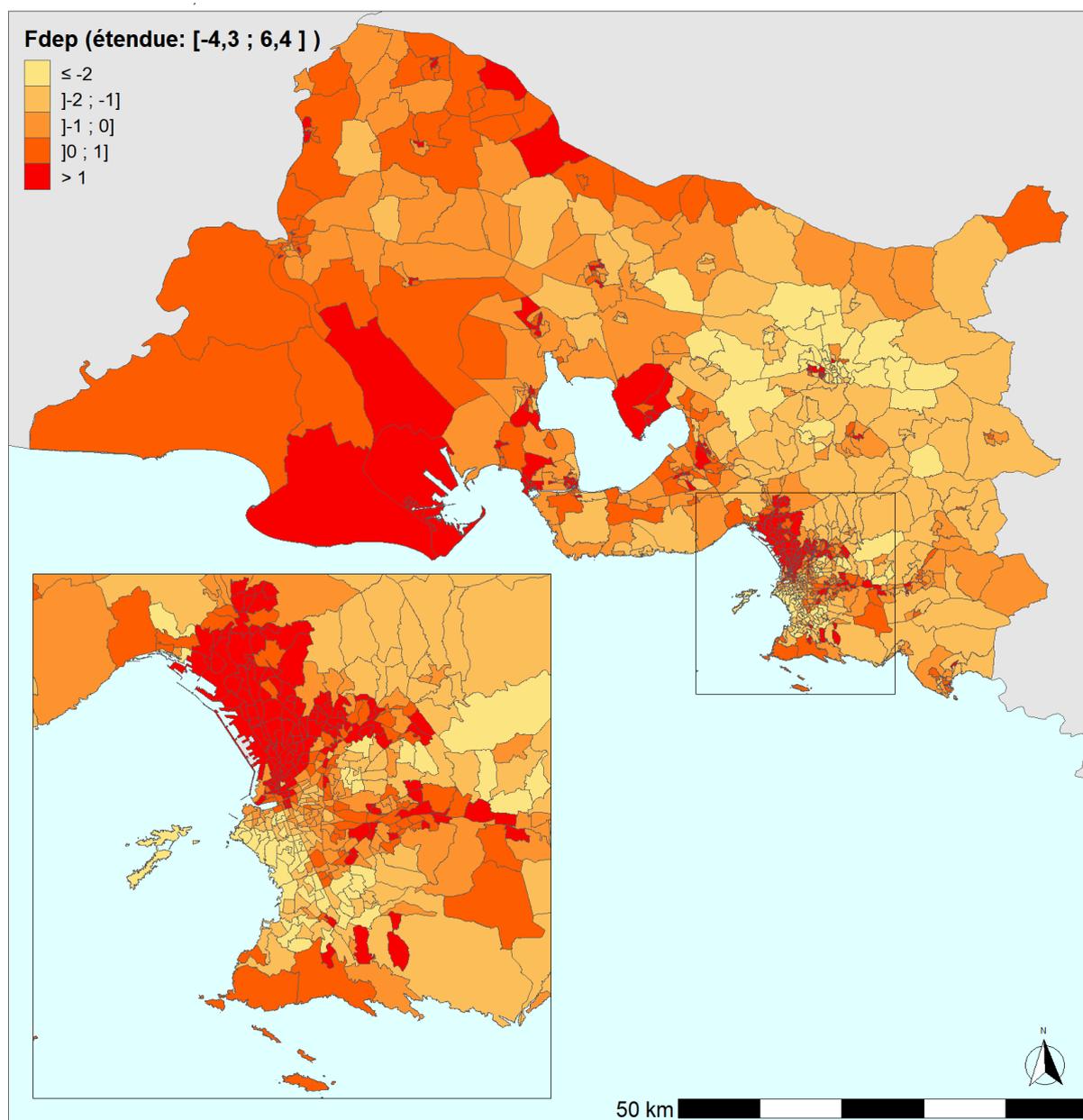
### 4a. Selon la commune ou l'arrondissement de résidence, Bouches-du-Rhône, 2015



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juin 2022

<sup>8</sup> <https://www.insee.fr/statistiques/5234149#titre-bloc-16>

#### 4b. Selon l'iris de résidence, Bouches-du-Rhône, 2015



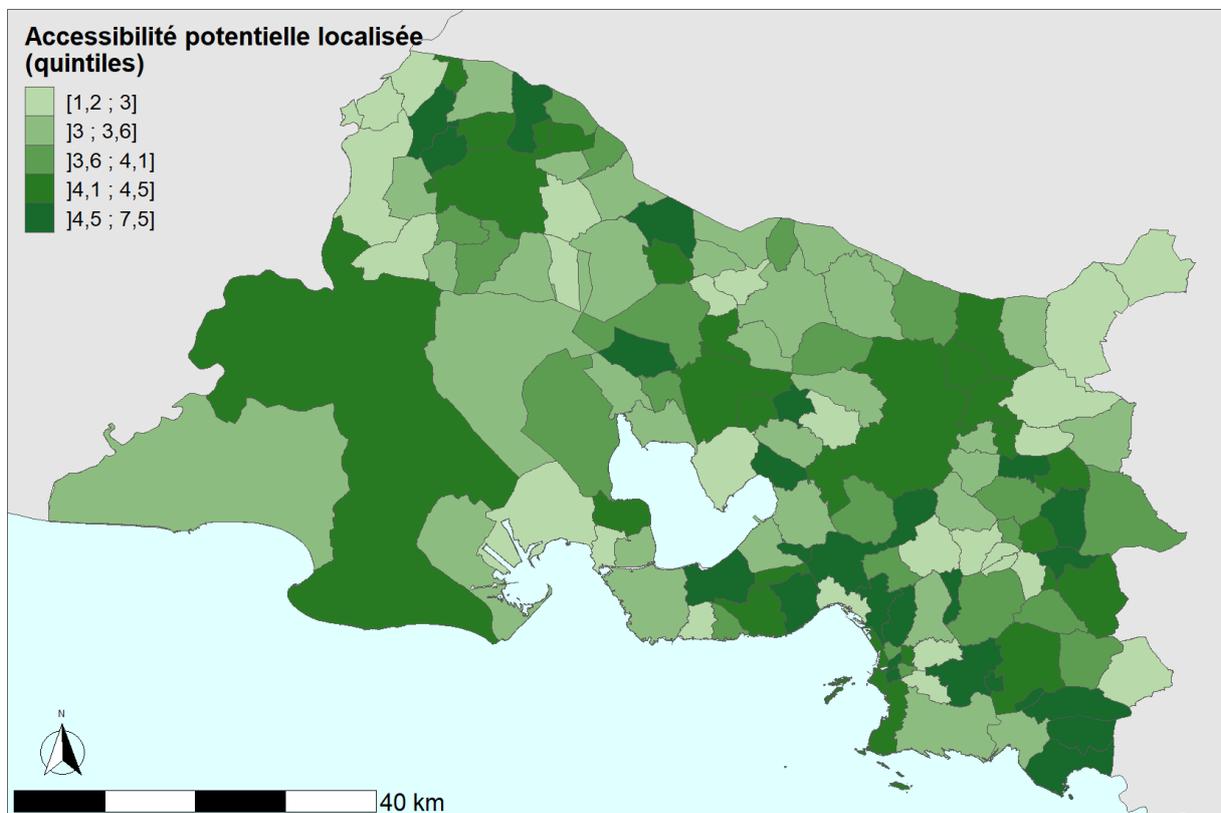
Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juin 2022

### 3.3 Offres de soins de premier recours

En 2017, le département des Bouches-du-Rhône présentait une densité relativement importante en médecins généralistes libéraux avec 112,5 médecins pour 100 000 habitants mais les médecins généralistes étaient inégalement répartis sur le territoire. Cette disparité était illustrée par la distribution de l'APL à l'échelle du département. Les communes les plus foncées disposaient des plus fortes densités de médecins et les communes les plus claires des plus faibles densités (Figure 5).

Les communes de Marseille et Aix-en-Provence concentrent plus de la moitié (55 %) de cette offre de soins alors que les zones de l'ouest et du nord du département, de même que les 10<sup>e</sup> et 15<sup>e</sup> arrondissements de Marseille présentent des densités inférieures ou égales à 80 médecins pour 100 000 habitants.

**Figure 5. Répartition (quintiles) de l'APL selon la commune de résidence, Bouches-du-Rhône, 2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juin 2022

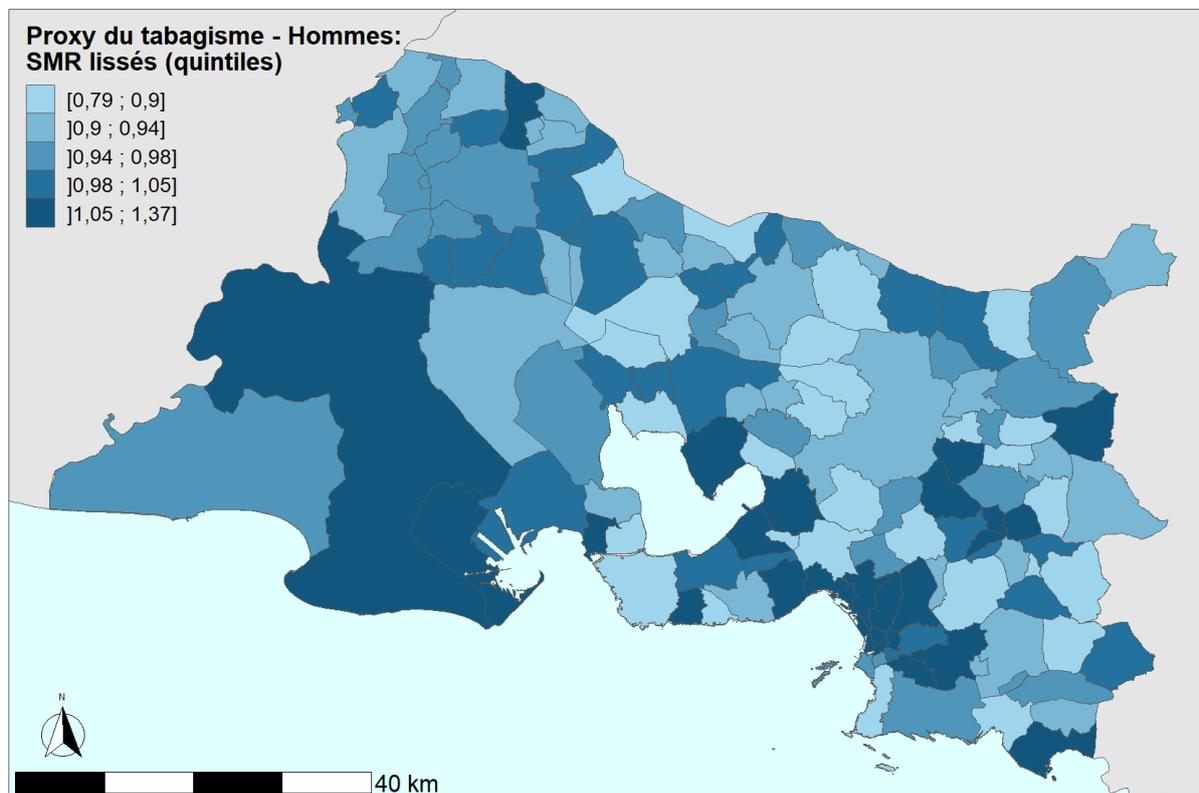
### 3.4 Proxy du tabagisme

Les données de mortalité par cancer du poumon ont été fournies par sexe. La distribution du *proxy* du tabagisme a été réalisée pour chaque sexe avec un dégradé de couleur, du plus clair pour les unités géographiques où le tabagisme est plus faible au plus foncé pour les unités géographiques où le tabagisme est plus important.

La distribution du *proxy* du tabagisme est différente selon le sexe. En effet, chez les hommes, les communes présentant une valeur élevée du *proxy* (représentant un tabagisme élevé) sont plutôt retrouvées dans les arrondissements du centre et du nord de Marseille ainsi que dans les communes situées à l'est et à l'ouest de l'étang-de-Berre (Figure 6a). Chez les femmes, l'ensemble des arrondissements de Marseille présente une valeur élevée du *proxy* du tabagisme ainsi que les communes situées au sud-est et nord-ouest du département (Figure 6b).

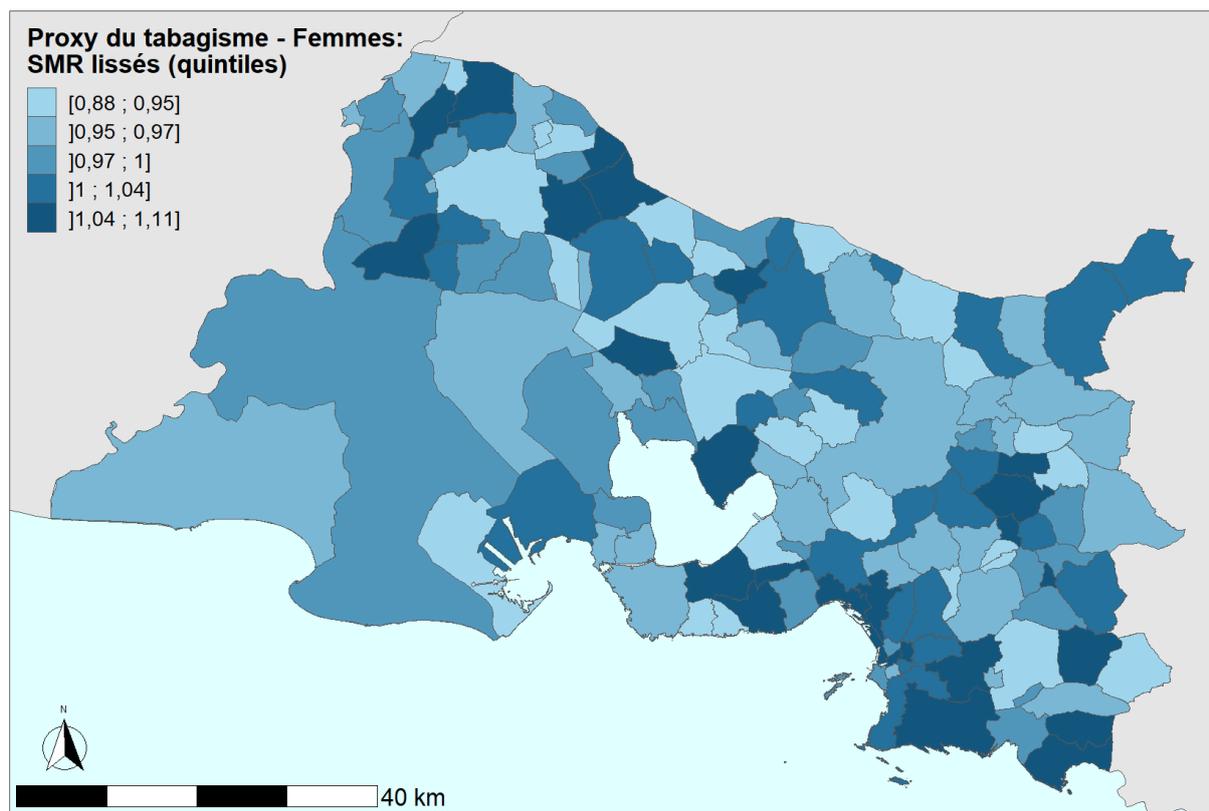
**Figure 6. Répartition par quintiles du *proxy* du tabagisme selon la commune de résidence, Bouches-du-Rhône, 2002-2012**

#### 6a. Chez les hommes



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - août 2022

## 6b. Chez les femmes



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - août 2022

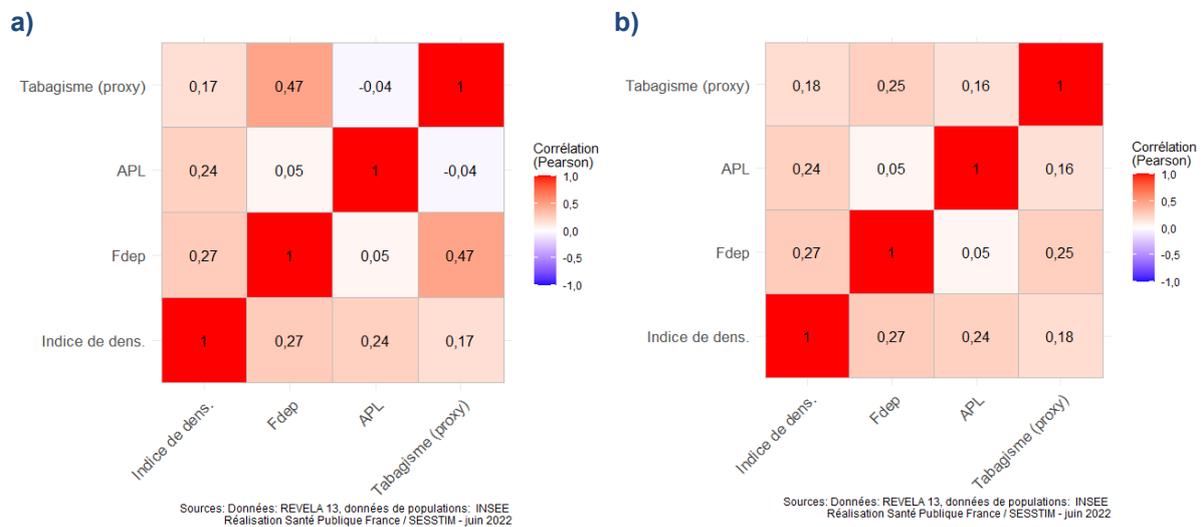
### 3.5 Corrélation des variables d'ajustement

L'ensemble des coefficients de corrélation de Pearson entre variables d'ajustement a été analysé à l'échelle communale pour les deux sexes (Figure 7).

L'indice de densité de l'Insee a été considéré comme une variable numérique de 1 à 4 (1 correspondant à la plus faible densité de population).

Les variables de confusion sont moyennement corrélées (< 0,5). La corrélation la plus élevée est retrouvée entre le *proxy* du tabagisme et la défaveur sociale (Fdep) chez les hommes. Cette interaction a été testée dans les modèles retenus pour les trois cancers chez les hommes.

**Figure 7. Matrices des coefficients de corrélation de Pearson des variables d'ajustement à l'échelle communale chez les hommes (a) et chez les femmes (b)**



## 4 RÉSULTATS

Les éléments concernant les caractéristiques cliniques, diagnostiques, épidémiologiques ainsi que les principaux facteurs de risques sont détaillés dans le premier rapport publié en 2019 et ne sont pas repris dans ce document. Ces éléments sont consultables sur la partie 3 du rapport précédent (6).

### 4.1 Caractéristiques sociodémographiques et histologiques des cas

#### 4.1.1 Cancers urothéliaux

Pour la période 2013-2018, 3 579 nouveaux cas de tumeurs malignes urothéliales, dont 2 896 chez les hommes et 683 chez les femmes, chez des adultes résidant dans le département des Bouches-du-Rhône au moment du diagnostic, sont présents dans la base de données de l'Observatoire REVELA13.

Le nombre moyen annuel de nouveaux cas de cancers urothéliaux est d'environ 596 cas dont 483 chez les hommes et 113 chez les femmes. Parmi ces cas, 3 169 (89 %) sont des cancers de la vessie, 187 (5 %) du bassinet, 181 (5 %) de l'uretère et 42 (1 %) des tumeurs malignes localisées sur un autre organe ou non précisé.

##### 4.1.1.1 Caractéristiques sociodémographiques

Le sexe ratio (H/F) est de 4,6 pour les cancers de vessie, 3 pour les tumeurs de l'uretère, 2,2 pour les tumeurs du bassinet et de 2,2 pour les autres localisations.

L'âge moyen des patients est significativement différent chez les hommes et chez les femmes pour les cancers de vessie et les cancers du bassinet. Les femmes sont diagnostiquées à un âge plus élevé que les hommes (76 ans en moyenne pour les deux cancers chez les femmes contre 74 ans pour la vessie et 71 ans pour le bassinet chez les hommes) (Tableau 1). La répartition par classe d'âge quinquennal et par sexe est présentée en Annexe 5.

La répartition des tumeurs urothéliales par catégorie d'âge est significativement différente en fonction du sexe ( $p < 0,001$ ). La classe d'âge des 65 à 84 ans représente 64 % des cas chez les hommes et 55 % chez les femmes. Celle des moins de 65 représente 20 % des cas parmi les deux sexes (Annexe 6).

Pour 74 %, les cas sont diagnostiqués dans une structure privée, 21 % dans une structure publique et 5 % au Centre de lutte contre le cancer (CLCC) de la région PACA.

Plus de la moitié des cas incidents recensés sont diagnostiqués à Marseille (60 %), 14 % dans le secteur d'Aix-en-Provence, 9 % dans le secteur d'Aubagne-La Ciotat, 7 % dans le secteur de Martigues, environ 3 % dans le secteur d'Arles et dans le secteur de Salon-de-Provence ainsi que dans les départements limitrophes.

**Tableau 1. Caractéristiques sociodémographiques des cancers urothéliaux selon la localisation tumorale et le sexe, Bouches-du-Rhône, 2013-2018**

Localisation tumorale			Hommes	Femmes	Total	p-valeur <sup>1</sup>
<b>Cancer de la vessie</b>	Effectifs		2 606	563	3 169	
	Âge au diagnostic (ans)	Moyenne	74,3	76,2	74,6	< 0,01
		Médiane	74,9	78,9	75,5	
		Étendue	18 - 100	37 - 102	18 - 102	
<b>Cancer du bassinot</b>	Effectifs		121	60	181	
	Âge au diagnostic (ans)	Moyenne	70,9	75,8	72,5	< 0,01
		Médiane	72,3	78,3	74,4	
		Étendue	32 - 93	47 - 92	32 - 94	
<b>Cancer de l'uretère</b>	Effectifs		140	47	187	
	Âge au diagnostic (ans)	Moyenne	71,6	72,5	71,8	0,61
		Médiane	73,5	73,6	73,6	
		Étendue	42 - 92	47 - 87	42 - 93	
<b>Cancer des organes urinaires, autres et non précisés</b>	Effectifs		29	13	42	
	Âge au diagnostic (ans)	Moyenne	71,5	71,7	71,6	0,96
		Médiane	73,3	71,9	72,6	
		Étendue	41 - 97	52 - 84	41 - 97	
<b>Total</b>	Effectifs		2 896	683	3 579	
	Âge au diagnostic (ans)	Moyenne	74,0	75,8	74,3	< 0,01
		Médiane	74,7	78,5	75,2	
		Étendue	18 - 100	37 - 102	18 - 102	

<sup>1</sup>Test de Student ou Wilcoxon - comparaison des âges moyens entre sexes

#### 4.1.1.2 Caractéristiques histologiques

La répartition des tumeurs urothéliales par stade anatomopathologique est significativement différente chez les hommes et chez les femmes ( $p < 0,001$ ) avec une proportion plus élevée de tumeurs envahissant le muscle ( $> pT1$ ) chez les femmes (51 %) que chez les hommes (41 %) (Tableau 2). Cette différence s'observe quelle que soit la classe d'âge considérée (Annexe 6). On note également une proportion plus importante de diagnostics suite à la découverte d'une métastase prévalente chez les femmes (2 %) que chez les hommes (1 %). En revanche, la sévérité de la maladie, évaluée par le grade de la tumeur, est globalement similaire pour les deux sexes.

**Tableau 2. Répartition du stade et du grade tumoral et des métastases prévalentes des tumeurs urothéliales, par sexe, Bouches-du-Rhône, 2013-2018**

	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 2 890	N = 678	N = 3 568	
<b>Stade tumoral</b>				<0,001
Stade pT1	1 684 (58 %)	321 (47 %)	2 005 (56 %)	
Stade > pT1	1 174 (41 %)	343 (51 %)	1 517 (43 %)	
Métastases prévalentes	32 (1 %)	14 (2 %)	46 (1 %)	
Manquant	6	5	11	
	N = 2 781	N = 637	N = 3 418	
<b>Grade tumoral</b>				0,052
Faible grade	47 (1 %)	19 (3 %)	66 (2 %)	
Bas grade	492 (18 %)	99 (16 %)	591 (17 %)	
Haut grade	2 242 (81 %)	519 (81 %)	2 761 (81 %)	
Manquant	115	46	161	

<sup>1</sup>Test du khi-deux d'indépendance

## 4.1.2 Cancers du rein

Pour la période 2013-2018, l'Observatoire REVELA13 comptabilise 1 991 cas incidents de cancers du parenchyme rénal chez des adultes, dont 1 366 hommes et 625 femmes, résidant dans le département des Bouches-du-Rhône au moment du diagnostic. Ils sont répartis en 1 943 de tumeurs malignes du rein avec un diagnostic histologique (98 %), dont 1 340 chez les hommes et 603 chez les femmes, et 48 tumeurs non confirmées par histologie (diagnostic par imagerie) (2 %). **Seuls les cas incidents avec diagnostic histologique ont été retenus dans la suite des analyses.**

Le nombre annuel moyen de cas incidents est de 324 cas dont 223 chez les hommes et 101 cas chez les femmes.

### 4.1.2.1 Caractéristiques sociodémographiques

Le sexe ratio (H/F) des cancers du rein est de 2,2. L'âge moyen et médian au moment du diagnostic est identique chez les hommes et les femmes (Tableau 3). La répartition par classe d'âge quinquennal et par sexe des cancers du rein est présentée en Annexe 5.

La répartition des cas par catégorie d'âge n'est pas différente en fonction du sexe contrairement aux données de la précédente analyse ( $p = 0,2$ ). La classe d'âge des 65 à 84 ans représente la moitié des cas pour les deux sexes et celle des moins de 65 ans, un peu moins de la moitié (48 %) (Annexe 7).

Un peu plus de la moitié des cas (56 %) est diagnostiquée dans une structure privée, 33 % dans une structure publique et 11 % au CLCC.

Comme pour les cancers urothéliaux, les cas sont principalement diagnostiqués à Marseille (61 %) et à Aix-en-Provence (16 %).

**Tableau 3. Caractéristiques sociodémographiques des cas de cancers du rein en fonction du sexe, Bouches-du-Rhône, 2013-2018**

Localisation tumorale			Hommes	Femmes	Total	p- valeur <sup>1</sup>
Cancers du rein	Effectifs		1 340	603	1 943	
	Âge au diagnostic	Moyenne	64,7	65,2	64,9	0,47
		Médiane	65,5	66,8	66,0	
	(ans)	Étendue	25 - 94	21 - 96	21 - 96	

<sup>1</sup>Test de Student - comparaison des âges moyens entre sexes

### 4.1.2.2 Caractéristiques histologiques

Parmi les 1 943 cas pour lesquels un grade tumoral était renseigné, la plupart sont de bas grades (grade 1 ou 2 de Führman). La répartition par grades tumoraux est similaire chez les hommes et les femmes (Tableau 4).

Parmi les 1 943 cas confirmés, 1 632 ont bénéficié d'une néphrectomie totale ou partielle permettant de disposer d'un stade tumoral. Deux tiers de ces patients sont de stade pT1. La répartition par stade des cancers du rein est similaire chez les hommes et chez les femmes (Tableau 4).

**Tableau 4. Répartition du stade et du grade tumoral et des métastases prévalentes des cancers du rein, par sexe, Bouches-du-Rhône, 2013-2018**

	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 1 175	N = 523	N = 1 698	
<b>Stade tumoral</b>				0,9
Stade pT1	763 (65 %)	334 (64 %)	1 097 (64 %)	
Stade > pT1	357 (30 %)	163 (31 %)	520 (31 %)	
Métastases prévalentes	55 (5 %)	26 (5 %)	81 (5 %)	
Manquant	165	80	245	
	N = 1 206	N = 527	N = 1 733	
<b>Grade tumoral</b>				0,2
Bas grade	804 (67 %)	367 (70 %)	1 171 (68 %)	
Haut grade	402 (33 %)	160 (30 %)	562 (32 %)	
Manquant	134	76	210	

<sup>1</sup>Test du khi-deux d'indépendance

Près des trois-quarts des cancers du parenchyme rénal sont des tumeurs à cellules claires chez les hommes et chez les femmes. La répartition des cas par type histologique des autres tumeurs est différente chez les hommes et chez les femmes (Tableau 5). Les tumeurs de type tubulo-papillaires sont un peu plus fréquentes chez les hommes alors que les tumeurs chromophobes sont plus fréquentes chez les femmes.

**Tableau 5. Répartition des cas de cancers du rein selon le type histologique en et du sexe, Bouches-du-Rhône, 2013-2018**

Histologie	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 1 323	N = 595	N = 1 918	
Cellules claires	958 (71 %)	443 (73 %)	1 400 (73 %)	
Tubulo-papillaires	252 (19 %)	77 (13 %)	329 (17 %)	
Chromophobe	74 (6 %)	53 (9 %)	127 (7 %)	
Sarcomatoïdes	4 (0 %)	2 (0 %)	6 (0 %)	0,001
Tubes collecteurs de Bellini	2 (0 %)	1 (0 %)	3 (0 %)	
Autre	33 (2 %)	20 (3 %)	53 (3 %)	
Manquant	17	8	25	

<sup>1</sup>Test exact de Fisher

### 4.1.3 Leucémies aiguës

Pour la période 2013-2018, 859 cas de leucémies aiguës, dont 457 chez les hommes et 402 chez les femmes, sont inclus dans la base de données de l'Observatoire REVELA13. Parmi ces cas, 127 sont des leucémies aiguës survenues après un syndrome myélodysplasique ou un syndrome myéloprolifératif déjà connu et 732 sont des leucémies aiguës incidentes, dont 378 chez les hommes et 354 chez les femmes. **Seules ces dernières ont été retenues pour calculer les incidences et réaliser les analyses.**

En moyenne, 122 cas de leucémies aiguës incidentes sont diagnostiqués chaque année sur le département dont 63 chez les hommes et 59 chez les femmes.

Parmi ces 732 cas, 599 correspondent à des leucémies aiguës myéloïdes (LAM) soit 82 % et 108 à des leucémies aiguës lymphoïdes / lymphome lymphoblastiques (LAL/LL) soit 15 %. Les autres cas sont répartis en 14 cas de lymphome de Burkitt, 4 cas de leucémies aiguës mixtes et 7 cas de leucémies aiguës inclassables.

Plus de la moitié des cas (59 %) est diagnostiquée à l'Institut Paoli Calmette (IPC), les autres étant principalement diagnostiqués par l'Assistance publique - Hôpitaux de Marseille (AP-HM)

(17 %) et le Centre hospitalier d'Aix-en-Provence (11 %). Les cliniques et hôpitaux privés des Bouches-du-Rhône sont à l'origine de 7 % de l'ensemble des diagnostics. L'hôpital d'Avignon prend en charge 5 % des diagnostics de leucémies aiguës de patients résidant sur le département des Bouches-du-Rhône.

#### 4.1.3.1 Caractéristiques sociodémographiques

Les caractéristiques des cas sont présentées uniquement pour les LAM et les LAL/LL, le nombre de cas étant trop faible pour les autres types.

Le sexe ratio H/F est de 1 pour les LAM et de 1,3 pour les LAL/LL. L'âge moyen n'est pas significativement différent en fonction du sexe que ce soit pour les LAM ou pour les LAL/LL, même si pour ces dernières, les femmes ont un âge moyen au diagnostic plus élevé (56,2 ans vs 49 ans pour les hommes). Tous sexes confondus, l'âge moyen est de 68 ans pour les LAM et de 52 ans pour les LAL/LL (Tableau 6). La répartition par classe d'âge quinquennal et par sexe des LAM et des LAL est présentée en Annexe 5.

La distribution des cas par classe d'âge est significativement différente entre les patients atteints d'une LAM et les patients atteints d'une LAL/LL. En effet, les trois-quarts des cas de LAM sont âgés de 59 ans et plus alors que la moitié des cas de LAL/LL est âgée de moins de 54 ans (Annexe 8).

Par ailleurs, pour les LAM, la répartition des cas par classe d'âge est significativement différente chez les hommes et les femmes ( $p < 0,05$ ). Pour les cas atteints de LAM, la proportion de cas âgés de plus de 85 ans est plus élevée chez les femmes (19 %) que chez les hommes (12 %) et la proportion de cas âgés de 65 à 84 ans est plus faible chez les femmes (43 %) que chez les hommes (53 %) (Annexe 8).

Pour les LAL/LL, la répartition est également différente mais non significative ( $p = 0,4$ ) potentiellement à cause des faibles effectifs. La proportion d'hommes de moins de 65 ans (73 %) est plus élevée que pour les femmes (60 %) (Annexe 8).

**Tableau 6. Caractéristiques sociodémographiques des cas de leucémies selon le type et le sexe, Bouches-du-Rhône, 2013-2018**

Type de LA			Hommes	Femmes	Total	p-valeur <sup>1</sup>
<b>LAM</b>	Effectifs		299	300	599	0,71
	Âge au diagnostic (ans)	Moyenne	67,8	68,3	68,0	
		Médiane	71,1	70,5	71,1	
		Étendue	19 - 94	19 - 97	19 - 97	
<b>LAL</b>	Effectifs		60	48	108	0,08
	Âge au diagnostic (ans)	Moyenne	49,0	56,2	52,2	
		Médiane	47,3	60,5	53,7	
		Étendue	18 - 94	18 - 92	18 - 94	
<b>Total</b>	Effectifs		359	348	707	0,71
	Âge au diagnostic (ans)	Moyenne	64,6	66,6	65,6	
		Médiane	68,0	69,2	68,8	
		Étendue	18 - 94	18 - 97	18 - 97	

<sup>1</sup>Test de Student - comparaison des âges moyens entre sexes

#### 4.1.3.2 Caractéristiques histologiques

##### A) Leucémies aiguës myéloïdes

Parmi les 599 LAM incidentes, 551 cas correspondent aux différents types de LAM de la classification de l'OMS et 48 cas n'entrent pas dans cette classification. Ces 551 cas se

répartissent en : 292 LAM sans spécificité (53 %), 110 LAM avec anomalies chromosomiques récurrentes (20 %), 99 LAM avec myélodysplasie (18 %), 46 LAM secondaire à traitement de cancer (8 %) et 4 LAM à cellules dendritiques (1 %). La classification n'est pas différente en fonction du sexe (Tableau 7).

**Tableau 7. Répartition des cas de LAM selon la classification OMS et le sexe, Bouches-du-Rhône, 2013-2018**

Type de LAM de la classification OMS	Hommes	Femmes	p-valeur <sup>1</sup>
	N = 279	N = 272	
LAM sans spécificité (LAM NOS)	152 (54 %)	140 (51 %)	0,09
LAM avec anomalies chromosomiques récurrentes (LAM ACR)	54 (19 %)	56 (21 %)	
LAM avec myélodysplasie (LAM MRC)	56 (20 %)	43 (16 %)	
LAM secondaire à un traitement de cancer	15 (5 %)	31 (11 %)	
LAM à cellules dendritiques	2 (1 %)	2 (1 %)	

<sup>1</sup> Test exact de Fisher

Les 599 cas de LAM incidentes ont également été classés selon la classification FAB lorsque celle-ci était précisée dans le compte rendu du myélogramme (Tableau 8). Les types les plus fréquents sont les LAM avec maturation (25 %), les leucémies aiguës myélomonocytaires (18 %) et monocytaires (19 %).

**Tableau 8. Répartition des cas de LAM selon la classification FAB et le sexe, Bouches-du-Rhône, 2013-2018**

Classification FAB	Hommes	Femmes	Total
	N = 232	N = 221	N = 453
Leucémie aiguë myéloïde peu différenciée (M0)	27 (12 %)	25 (11 %)	52 (11 %)
Leucémie aiguë myéloïde sans maturation (M1)	25 (11 %)	40 (18 %)	65 (14 %)
Leucémie aiguë myéloïde avec maturation (M2)	64 (28 %)	51 (23 %)	115 (25 %)
Leucémie aiguë promyélocytaire (M3)	16 (6,9 %)	18 (8,1 %)	34 (7,5 %)
Leucémie aiguë myélomonocytaire (M4)	43 (19 %)	37 (17 %)	80 (18 %)
Leucémie aiguë monocytaire (M5)	44 (19 %)	42 (19 %)	86 (19 %)
Leucémie aiguë érythrocytaire ou érythroïde (M6)	12 (5,2 %)	5 (2,3 %)	17 (3,8 %)
Leucémie aiguë mégacaryoblastique (M7)	1 (0,4 %)	3 (1,4 %)	4 (0,9 %)
Non classé	67	79	146

### B) Leucémies aiguës lymphoïdes

Parmi les 108 cas de LAL/LL, les LAL/LL B sont majoritaires (75 %), dont 72 % sont des LAL communes (LALB2). Les LAL/LL de lignée T représentent 19 % des cas, et les leucémies aiguës ou lymphome à cellules de Burkitt, 6 % des cas (Tableau 9).

**Tableau 9. Répartition des cas de LAL selon leur classification, Bouches-du-Rhône, 2013-2018**

Type de leucémie : leucémie aiguë lymphoïde / lymphome lymphoblastique (LAL/LL)	Total
	N = 108
Leucémie aiguë / lymphome à cellules de Burkitt /Burkitt like	7 (6 %)
LAL/ LL de la lignée B	60 (56 %)
LAL/ LL de la lignée B avec chromosome Philadelphia	21 (19 %)
LAL/ LL de la lignée T	20 (19 %)

Une description détaillée des leucémies aiguës selon la classification CIMO3 est présentée en Annexe 9.

## 4.2 Incidence départementale et répartition spatiale des cas

Pour les analyses spatiales, à l'échelle des communes ou des Iris, et ce, pour l'ensemble des trois cancers et pour les deux sexes, les valeurs de DIC des modèles complets (ajustés sur la variable de densité de population, l'indice de défaveur sociale, l'APL et le *proxy* du tabagisme) diffèrent de moins de 10 points de la valeur de DIC la plus faible. De plus, les résultats de lissages des SIR sont globalement stables selon le choix du modèle.

**Les modèles complets ont donc été retenus dans cette étude afin de pouvoir interpréter de façon comparative les effets des facteurs d'ajustement par cancer et par sexe.**

L'Annexe 3 présente en détail les différentes étapes de la modélisation : le choix des paramètres, l'analyse de la convergence des estimations *a posteriori*, ainsi que la comparaison des DIC des modèles.

### 4.2.1 Cancers de la vessie

#### 4.2.1.1 Incidence départementale estimée

L'incidence départementale du cancer de la vessie standardisée monde tous âges confondus est estimée à 18,8 cas pour 100 000 personnes-années chez les hommes et à 3,0 cas pour 100 000 personnes-années chez les femmes, soit des taux d'incidence significativement supérieurs à ceux estimés au niveau national sur la même période (Tableau 10).

**Tableau 10. Taux d'incidence bruts et standardisés monde (TSM) pour cancer de vessie pour 100 000 personnes-années, Bouches-du-Rhône 2013-2018 (incluant les cas de moins de 20 ans)**

Sexe	Bouches-du-Rhône 2013-2018			France 2013-2018	
	Nombre de cas	Taux brut	TSM [IC 95 %]	Taux brut	TSM [IC 95 %]
Femmes	563	8,9	3,0 [2,7 ; 3,3]	7,1	2,3 [2,3 ; 2,4]
Hommes	2 606	45,0	18,8 [18,0 ; 19,6]	32,6	14,4 [14,3 ; 14,5]

#### 4.2.1.2 Répartition spatiale des risques

##### A) Résultats chez les hommes

Douze communes rurales ne comptent aucun cas de cancer de la vessie sur la période 2013-2018, ces communes représentent 1 % de la population totale masculine du département. Quatre-vingt-douze Iris (11 %) ne comptent aucun cas de cancer de la vessie sur la période 2013-2018, ces Iris représentent 7 % de la population totale masculine du département.

##### Cartographie des risques lissés à l'échelle de la commune

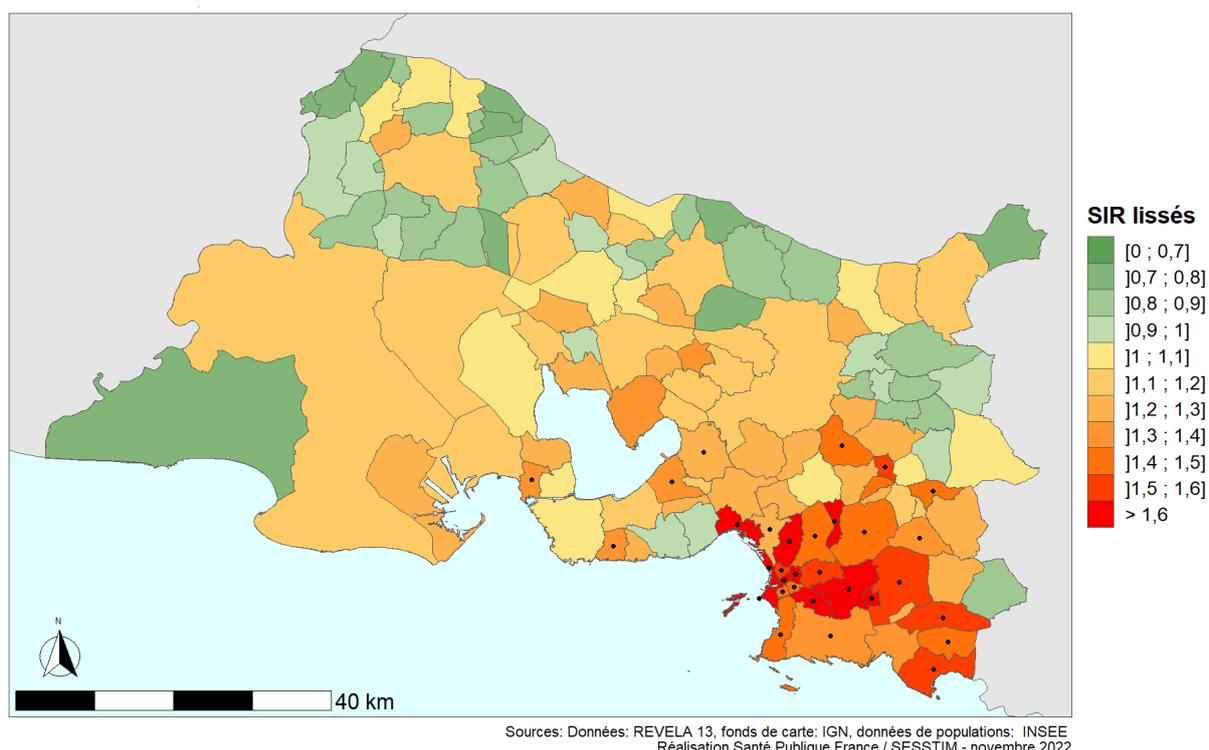
Les résultats du modèle ajusté sur l'ensemble des facteurs montrent une association positive significative importante avec le tabagisme (RR = 2,17 IC95 % [1,20 ; 3,95]). La densité de population est également positivement associée à la valeur du SIR, pour les communes de densité intermédiaire (RR = 1,39 IC95 % [1,14 ; 1,72]) et pour les communes densément peuplées (RR = 1,72 IC95 % [1,37 ; 2,16]) par rapport aux communes de densité faible (Annexe 10).

L'absence d'autocorrélation des résidus est validée grâce à l'indice de Moran ( $p = 1,00$ ).

L'analyse de sensibilité intégrant au modèle l'interaction entre le *proxy* du tabagisme et l'indice Fdep chez les hommes n'a pas fourni de résultat significatif, rejetant son intégration au modèle.

Les SIR lissés varient de 0,76 à 1,84 en médiane *a posteriori* avec une médiane de 1,13. Ils sont globalement supérieurs à 1 sur une grande partie du département (Figure 8). Une surincidence par rapport au niveau national est observée dans 31 communes : 27 situées dans la partie sud-est du département ainsi que les communes de Port-de-Bouc, Marignane, Vitrolles et Sausset-les-Pins, situées autour ou au sud de l'étang de Berre. La zone continue au sud-est inclut l'ensemble des arrondissements de Marseille et 12 communes situées au nord et à l'est de Marseille, jusqu'à la frontière avec le département du Var, avec au sud les communes d'Aubagne, Ceyreste ou La Ciotat, et au nord les communes de Gardanne, Gréasque et La Bouilladisse. Aucune commune ne présente de sous-incidence sur le département.

**Figure 8. Cartographie des SIR lissés de cancers de la vessie chez les hommes selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



L'analyse de sensibilité prenant en compte les informations disponibles des cas de cancer de la vessie pris en charge par les urologues ayant refusé de participer au recueil de données ne modifie pas les résultats précédents (légère augmentation des valeurs de SIR dans les communes de Port-de-Bouc et Martigues) (Annexe 11).

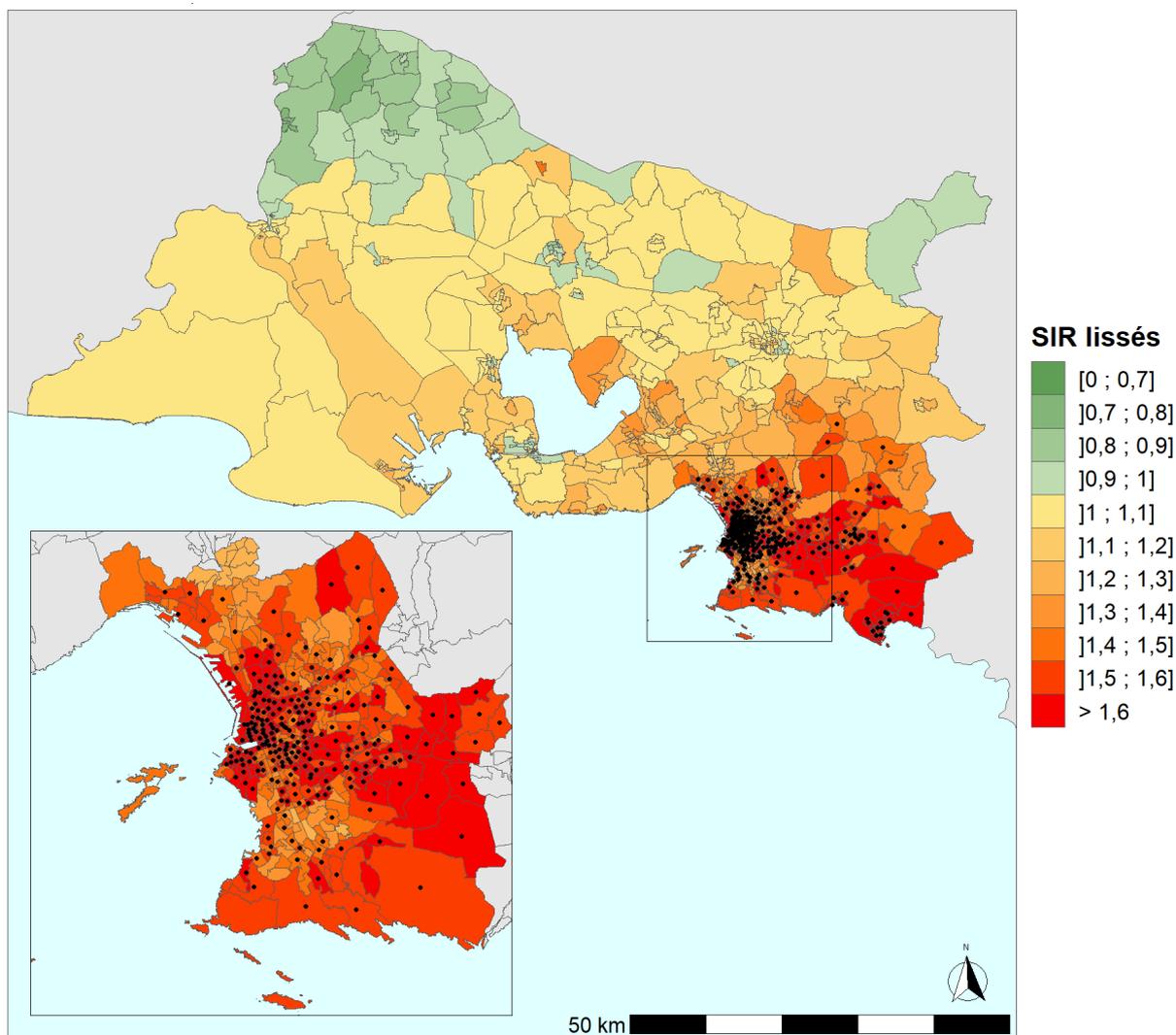
#### *Cartographie des risques lissés à l'échelle des Iris*

À l'échelle de l'Iris, les résultats du modèle de lissage BYM ajusté sur l'ensemble des facteurs d'ajustement ne montrent plus d'association significative entre les SIR et les covariables. Cependant, l'estimation ponctuelle de l'effet du tabagisme est toujours élevée à cette échelle (RR = 1,45 IC95 % [0,92 ; 2,29]) contrairement à celles liées à la densité de population (Annexe 10).

À l'échelle de l'Iris, les SIR lissés varient de 0,76 à 2,05 en médiane *a posteriori* avec une médiane à 1,38. Ils sont globalement supérieurs à 1 sur l'ensemble du département excepté la pointe nord-ouest du département. Une surincidence significative de valeurs médianes supérieures à 1,6 est observée en continu depuis l'est de Marseille jusqu'au département du Var, couvrant les communes de Cassis et La Ciotat et s'étendant au nord sur une partie

importante d'Aubagne jusqu'à Cuges-les-Pins. Au nord de cette zone, des surincidences significatives sont observées en continu jusqu'à la commune de Gardanne (Figure 9). Aucun Iris ne présente de sous-incidence significative sur le département.

**Figure 9. Cartographie des SIR lissés de cancers de la vessie chez les hommes selon l'Iris de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

### Identification de clusters à l'échelle de la commune

Les méthodes globales de clustering montrent une hétérogénéité spatiale liée à une surdispersion des SIR de cancers de la vessie (Potthoff-Whittinghill ;  $p < 10^{-4}$ ), une absence de corrélation spatiale globale (Moran et EBI : NS) et l'existence de clusters dans la zone d'étude (Tango ;  $p < 10^{-4}$ ).

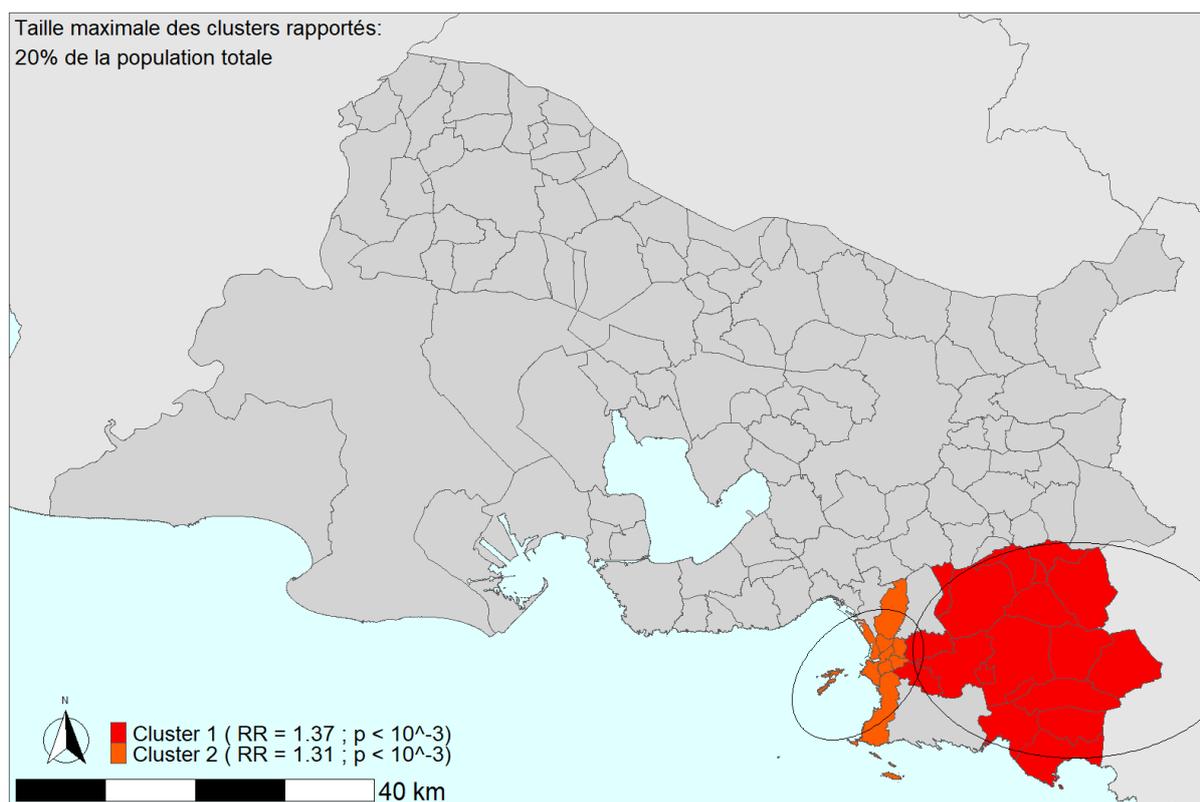
La Figure 10 présente les résultats de la recherche de clusters significatifs de surincidence par SaTScan. Le jeu de clusters optimal sélectionné par l'indice de Gini est retrouvé pour une taille de fenêtre de 20 %.

Deux clusters ont été identifiés :

- le premier inclut les communes de la partie sud-est du département, Allauch, Aubagne, Auriol, Cadolive, Carnoux-en-Provence, Cassis, Ceyreste, Cuges-les-Pins, Gémenos,

- La Bouilladisse, La Ciotat, La Destrousse, La Penne-sur-Huveaune, Peypin, Plan-de-Cuques, Roquefort-la-Bédoule, Roquevaire ainsi que les 10<sup>e</sup>, 11<sup>e</sup> et 12<sup>e</sup> arrondissements de Marseille avec 639 cas observés pour 518 cas attendus et un RR de 1,37 par rapport au reste des communes (1 950 cas observés et 2 092 cas attendus).
- le second cluster concerne 9 arrondissements de Marseille, du 1<sup>er</sup> au 8<sup>e</sup> ainsi que le 14<sup>e</sup> avec 654 cas observés pour 512 cas attendus et un RR de 1,31 par rapport au reste des communes (1 965 cas observés et 2 086 cas attendus).

**Figure 10. Représentation des clusters significatifs de cancers de la vessie chez les hommes à l'échelle de la commune, Bouches-du-Rhône, 2013-2018**

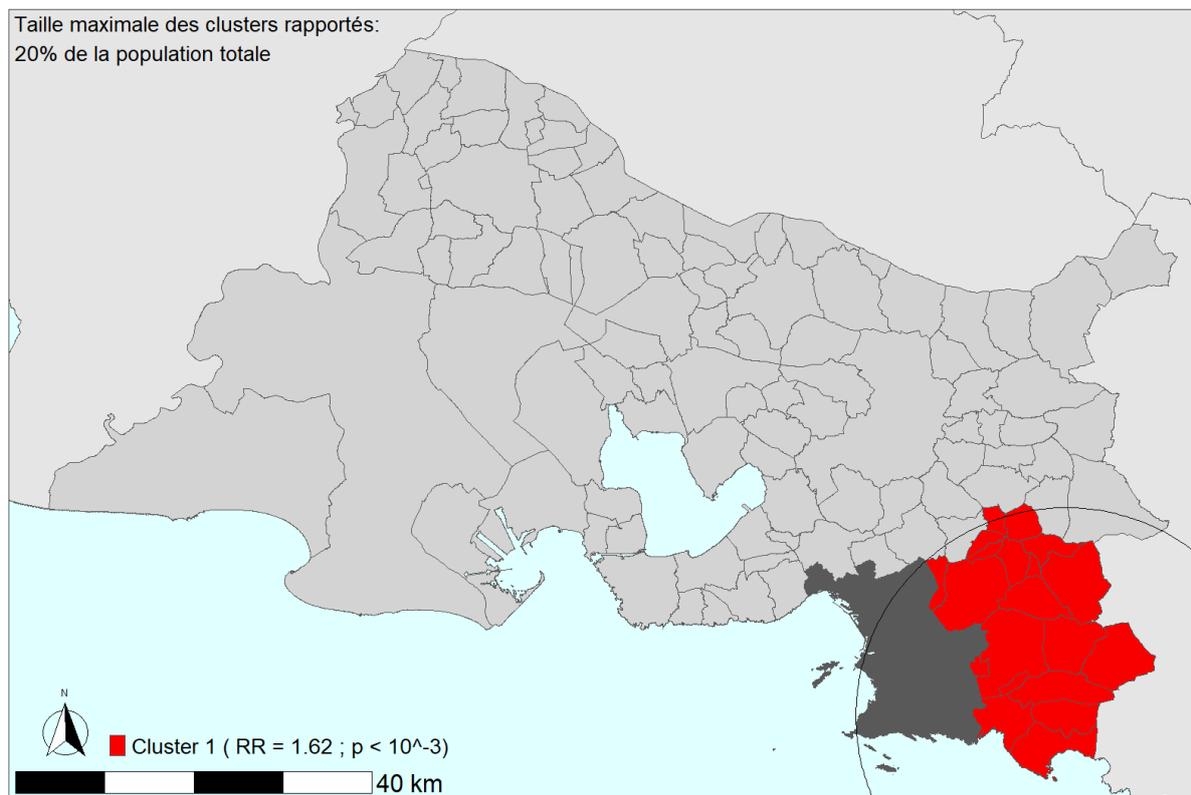


Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juillet 2022

La commune de Marseille représente plus de 40 % de la population totale du département, et 46 % des cas incidents de cancer de vessie diagnostiqués chez l'homme sur le département des Bouches-du-Rhône entre 2013 et 2018, résidaient dans cette commune. Une recherche de clusters a donc été réalisée sur le territoire après exclusion de la commune de Marseille, pour contrôler le poids de la ville par rapport au reste du département.

Seul le cluster regroupant les mêmes communes du sud-est du département, ainsi que les communes de Belcodène, Gréasque et Saint-Savournin à sa frontière nord, est retrouvé avec 387 cas observés pour 267 cas attendus et un RR de 1,62 par rapport au reste des communes (1 003 cas observés et 1 223 cas attendus). L'exclusion de la commune de Marseille ne fait pas apparaître de nouveaux clusters dans aucune autre partie du territoire (Figure 11).

**Figure 11. Représentation des clusters significatifs de cancers de la vessie chez les hommes à l'échelle de la commune en excluant Marseille, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juillet 2022

### Identification de clusters à l'échelle de l'Iris

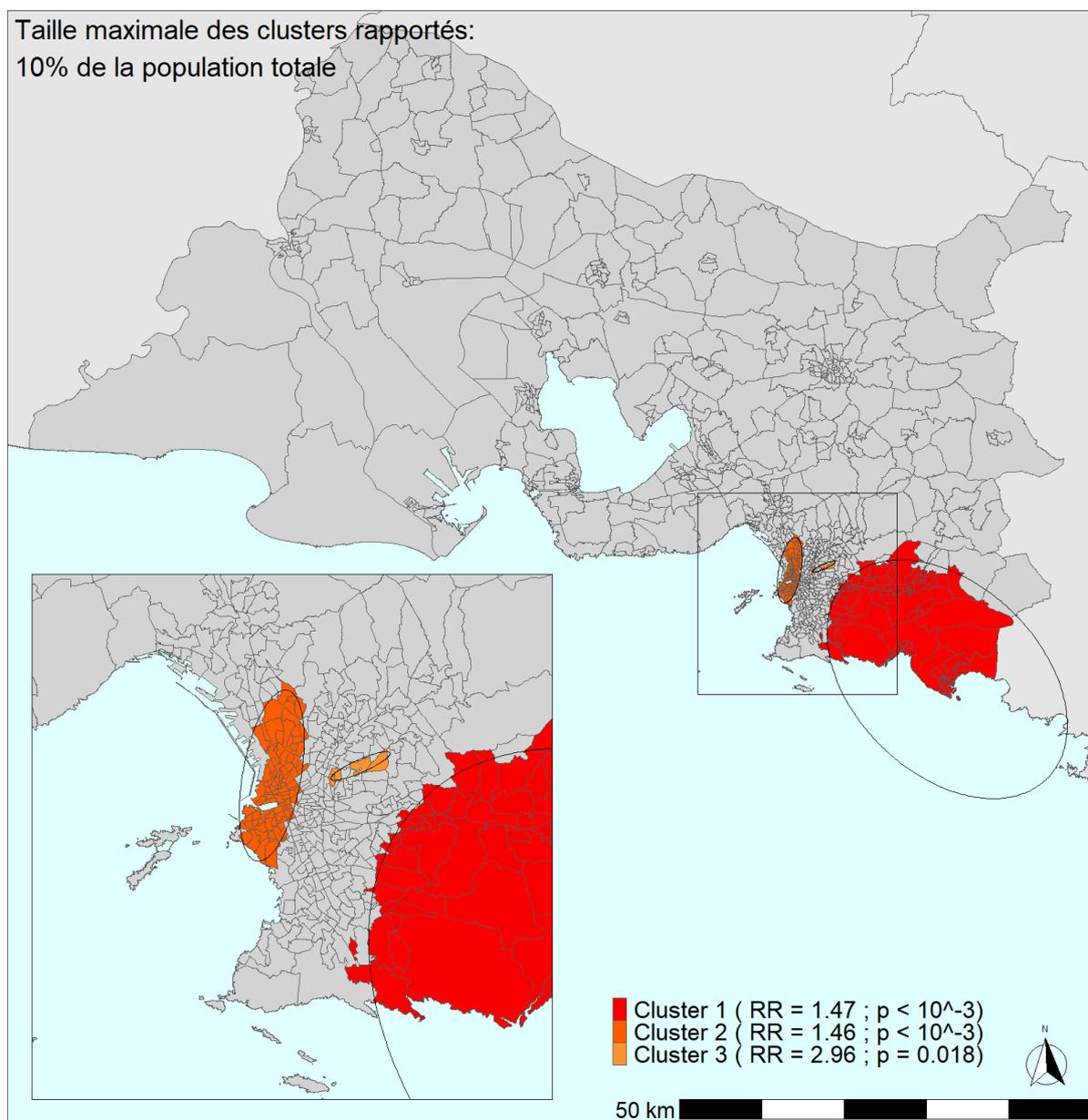
Les méthodes globales de clustering montrent une hétérogénéité spatiale liée à une surdispersion des SIR de cancers de la vessie (Potthoff-Whittinghill ;  $p < 10^{-4}$ ), la présence d'une corrélation spatiale globale (Moran et EBI :  $p < 10^{-4}$ ) et l'existence de clusters dans la zone d'étude (Tango ;  $p < 10^{-4}$ ).

La Figure 12 présente les résultats à l'échelle des Iris avec une taille maximale de la fenêtre de détection de 10 % de la population totale du département.

Trois clusters ont été identifiés correspondant à des zones restreintes des clusters identifiés à la commune :

- le premier inclut la partie sud du cluster détecté dans la même zone à l'échelle communale, se recentrant sur les communes de Cassis et La Ciotat, avec 318 cas observés pour 215 cas attendus et un RR à 1,47 par rapport au reste des iris (2 286 cas observés et 2 378 cas attendus). Les communes concernées étaient Aubagne, Carnoux-en-Provence, Cassis, Ceyreste, Gémenos, La Ciotat, La Penne-sur-Huveaune et Roquefort-la-Bédoule, ainsi que les 9<sup>e</sup>, 10<sup>e</sup> et 11<sup>e</sup> arrondissements de Marseille.
- le second, également détecté dans la commune de Marseille, inclut les 1<sup>er</sup>, 2<sup>e</sup>, 3<sup>e</sup>, 6<sup>e</sup>, 7<sup>e</sup> et une partie des 14<sup>e</sup> et 15<sup>e</sup> arrondissements de Marseille avec 318 cas observés pour 215 cas attendus et un RR à 1,46 par rapport au reste des iris (2 305 cas observés et 2 392 cas attendus).
- le 3<sup>e</sup> cluster ne comprend que quatre Iris des 4<sup>e</sup> et 12<sup>e</sup> arrondissement avec 36 cas observés pour 12 cas attendus et un RR à 2,96 par rapport au reste des iris (2 568 cas observés et 2 592 cas attendus).

**Figure 12. Représentation des clusters significatifs de cancers de la vessie chez les hommes à l'échelle de l'Iris, Bouches-du-Rhône, 2013-2018**

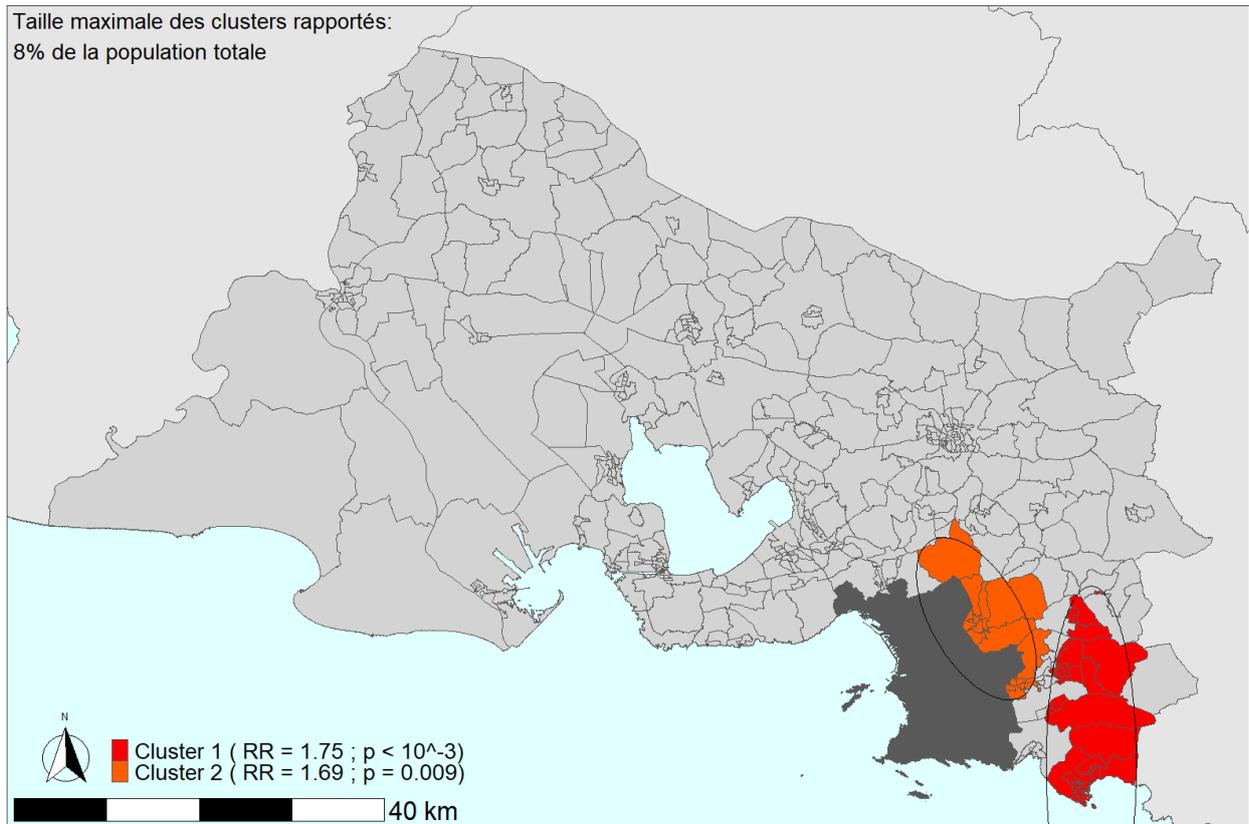


Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juillet 2022

L'exclusion de la ville de Marseille recentre les clusters identifiés au sud-est du département pour le premier et au nord-est de Marseille pour le deuxième (Figure 13) :

- le premier cluster apparaît sur une verticale allant de La Ciotat à Roquevaire et inclut des Iris des communes d'Aubagne, Auriol, Carnoux-en-Provence, Ceyreste, Géménos et Roquefort-la-Bédoule avec 183 cas observés pour 111 cas attendus et un RR de 1,75 par rapport au reste des iris (1 207 cas observés et 1 279 cas attendus).
- le second cluster inclut des Iris des communes d'Allauch, Aubagne, La Penne-sur-Huveaune, Plan-de-Cuques et Simiane-Collongue avec 106 cas observés pour 65 cas attendus et un RR de 1,69 par rapport au reste des iris (1 284 cas observés et 1 325 cas attendus).

**Figure 13. Représentation des clusters significatifs de cancers de la vessie chez les hommes à l'échelle de l'Iris en excluant la commune de Marseille, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juillet 2022

## B) Résultats chez les femmes

Quarante et une communes, représentant 7 % de la population féminine du département, ne recensent aucun cas de cancer de vessie pour la période 2013-2018.

Quatre cent-quarante-huit Iris (55 %) ne comptent aucun cas de cancer de la vessie sur la période 2013-2018, ces Iris représentent 50 % de la population totale féminine du département. Compte tenu du grand nombre d'Iris sans aucun cas de cancer de vessie recensé, les analyses spatiales n'ont pas été réalisées à cette échelle.

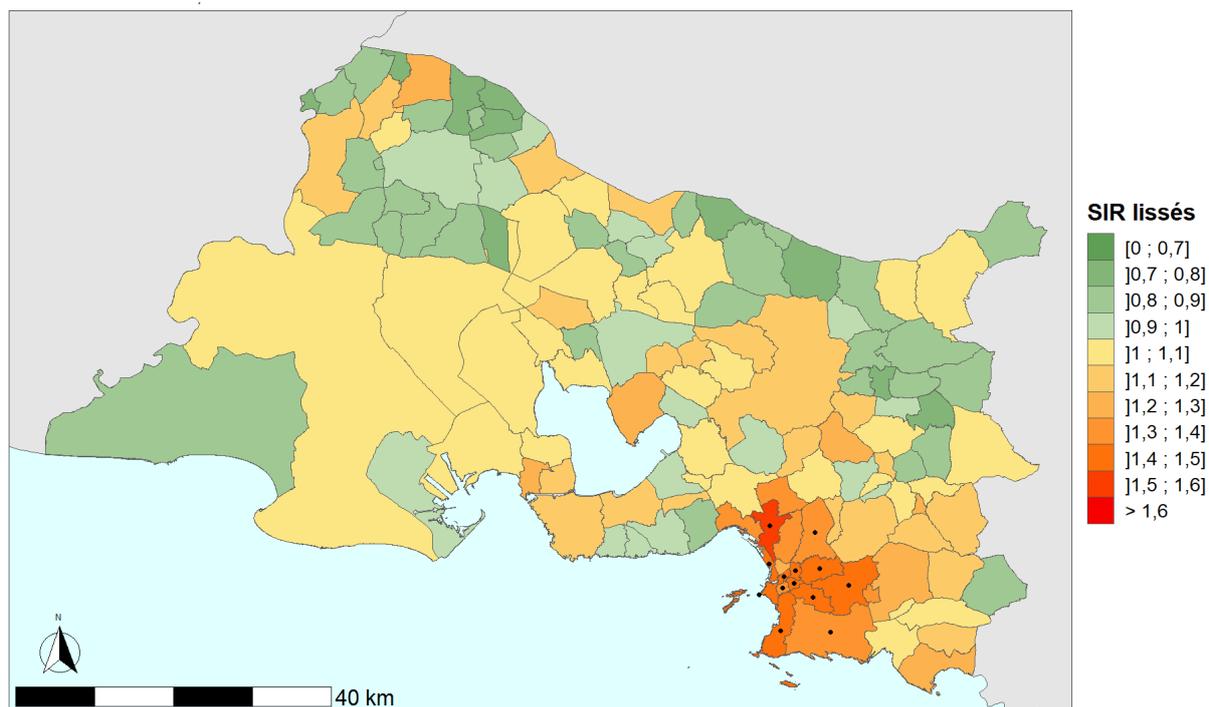
### Cartographie des risques lissés à l'échelle de la commune

Les résultats du modèle ajusté sur l'ensemble des covariables ne montrent aucune association significative entre les SIR et les covariables. Les communes les plus densément peuplées semblent néanmoins présenter un risque plus élevé que les communes de densité faible (RR = 1,57 IC95 % [0,96 ; 2,42]). Un effet positif non significatif du *proxy* du tabagisme est également observé (RR = 2,13 IC95 % [0,56 ; 8,64] (Annexe 10).

L'absence d'autocorrélation des résidus est validée grâce à l'indice de Moran (p = 0,62).

Les SIR lissés varient de 0,75 à 1,54 en médiane *a posteriori* avec une médiane de 1,04. Une surincidence par rapport à la moyenne nationale est observée pour l'ensemble des arrondissements de la ville de Marseille à l'exclusion des 3<sup>e</sup>, 14<sup>e</sup> et 16<sup>e</sup> arrondissements. Aucune commune ne présente de sous-incidence sur le département (Figure 14).

**Figure 14. Cartographie des SIR lissés de cancers de la vessie chez les femmes selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



L'analyse de sensibilité prenant en compte les informations disponibles sur les cas de cancer de la vessie pris en charge par les urologues ayant refusé de participer au recueil de données ne modifie pas les résultats précédents (légère augmentation des valeurs de SIR dans les communes concernées) (Annexe 11).

#### *Identification de clusters*

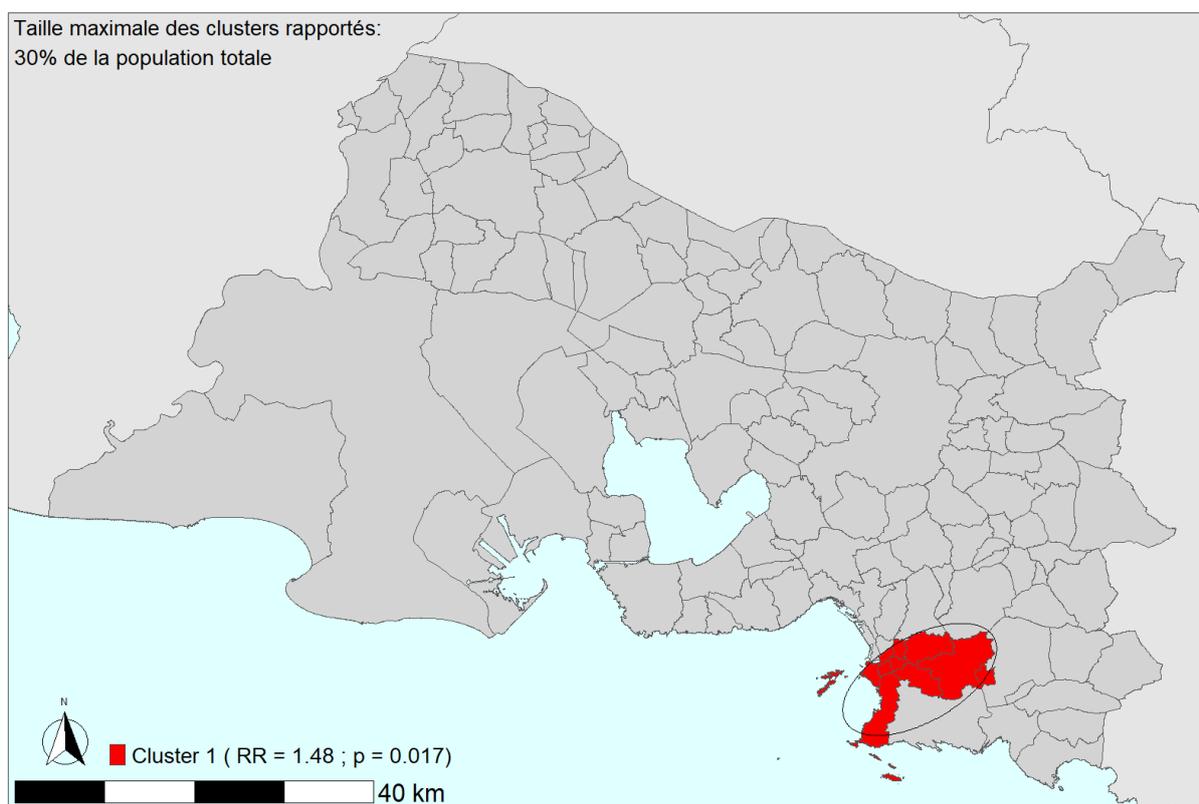
Les méthodes de clustering montrent une hétérogénéité spatiale liée à une surdispersion des SIR de cancers de la vessie (Potthoff-Whittinghill ;  $p < 10^{-4}$ ), l'absence d'une corrélation spatiale globale (Moran et EBI : NS) et l'existence de clusters dans la zone d'étude (Tango ;  $p = 0,04$ ).

La statistique de SaTScan sélectionne un jeu de clusters optimal correspondant à une taille de fenêtre rapportée de 30 %.

L'unique cluster détecté sur le département à l'échelle de la commune regroupe 8 arrondissements de Marseille (les 1<sup>er</sup>, 4<sup>e</sup>, 5<sup>e</sup>, 6<sup>e</sup>, 7<sup>e</sup>, 8<sup>e</sup>, 10<sup>e</sup> et 11<sup>e</sup>) et la commune de La Penne-sur-Huveaune avec 193 cas observés pour 146 cas attendus et un RR à 1,48 par rapport au reste des communes (370 cas observés et 417 cas attendus) (Figure 15).

La commune de Marseille représentait 50 % des cas incidents de cancer de vessie recensés chez les femmes sur le département des Bouches-du-Rhône entre 2013 et 2018. La recherche de clusters réalisée sur le département, après exclusion de la commune de Marseille, n'identifie plus de cluster significatif (non représenté).

**Figure 15. Représentation des clusters significatifs de cancers de la vessie chez les femmes à l'échelle de la commune, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juillet 2022

## 4.2.2 Cancers du rein

### 4.2.2.1 Incidence départementale estimée

L'incidence départementale du cancer du rein (regroupant les cancers du parenchyme rénal et les tumeurs urothéliales du bassin, de l'uretère et autres) standardisée monde tous âges confondus est estimée à 15,7 cas pour 100 000 personnes-années chez les hommes et à 6,2 cas pour 100 000 personnes-années chez les femmes. Elle est du même ordre de grandeur que celui observé en métropole pour les hommes, et légèrement inférieur à celui de la France pour les femmes, à la limite de la sous-incidence (Tableau 11).

**Tableau 11. Taux d'incidence bruts et standardisés monde (TSM) pour cancer du rein pour 100 000 personnes-années, Bouches-du-Rhône 2013-2018 (incluant les cas de moins de 20 ans)**

Sexe	Bouches-du-Rhône 2013-2018			France 2013-2018	
	Nombre de cas	Taux brut	TSM [IC 95 %]	Taux brut	TSM [IC 95 %]
Femmes	733	11,6	6,2 [5,7 ; 6,7]	14,2	6,8 [6,7 ; 6,9]
Hommes	1 635	28,2	15,7 [14,9 ; 16,6]	30,3	16,4 [16,2 ; 16,5]

#### 4.2.2.2 Répartition spatiale des risques

##### A) Résultats chez les hommes

Dix communes, représentant 1 % de la population masculine du département, ne comptent aucun cas de cancer du rein recensé entre 2013 et 2018.

Cent soixante-cinq Iris, représentant plus de 20 % des Iris, ne comptabilisent aucun cas de cancer du rein chez l'homme sur la période 2013-2018, ces Iris représentent 15 % de la population totale masculine du département. Compte tenu du grand nombre d'Iris sans aucun cas de cancer du rein recensé, les analyses spatiales n'ont pas été réalisées à cette échelle.

##### Cartographie des risques lissés à l'échelle de la commune

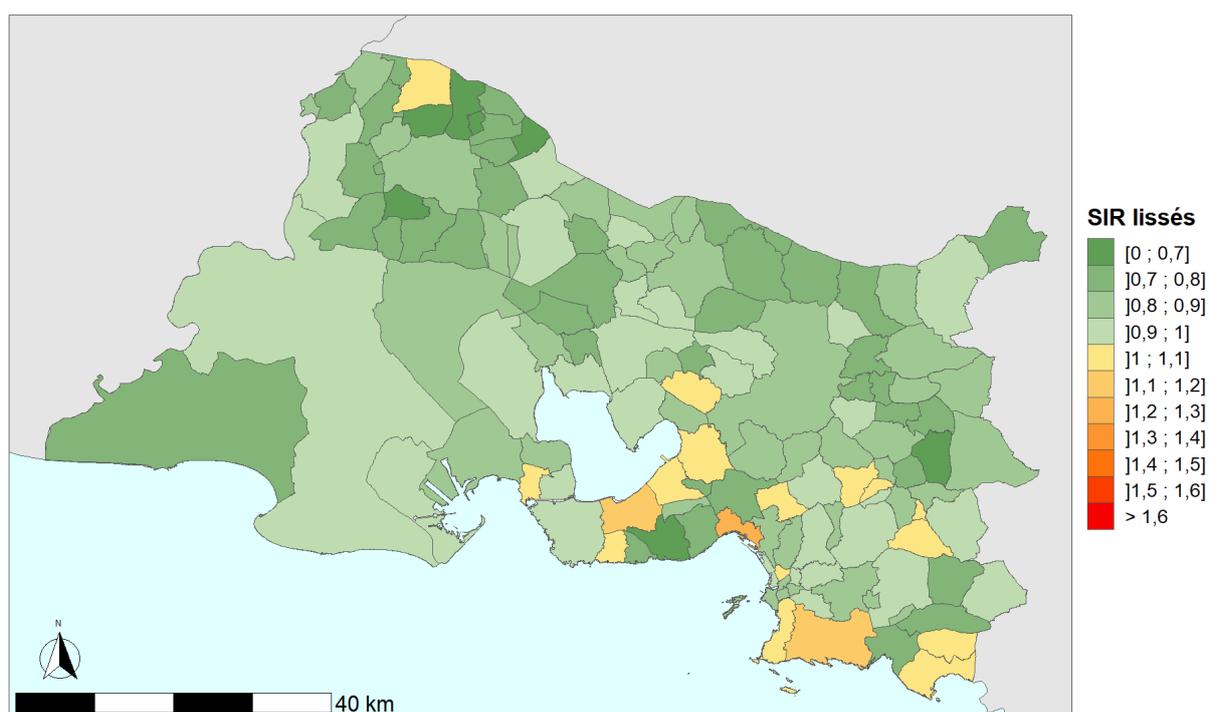
Les résultats du modèle de lissage ajusté sur l'ensemble des covariables ne montrent aucune association significative entre les SIR et les covariables. L'indice de densité de population de l'Insee et le *proxy* du tabagisme montrent néanmoins des valeurs d'effets positifs assez importants (Annexe 10).

L'absence d'autocorrélation des résidus est confirmée par un indice de Moran faible ( $\rho = 0,82$ ).

L'analyse de sensibilité intégrant au modèle l'interaction entre le *proxy* du tabagisme et l'indice Fdep chez les hommes n'a pas fourni de résultat significatif, rejetant son intégration au modèle.

Les SIR lissés varient de 0,63 à 1,21 en médiane *a posteriori* avec une médiane de 0,84. Ils sont globalement inférieurs à 1 sur le département, avec quelques valeurs supérieures à 1 principalement regroupées dans le quart sud-est (Figure 16). Aucune commune ne présente de surincidence significative. Dix-huit communes présentent une sous-incidence significative sur le département par rapport au niveau national.

**Figure 16. Cartographie des SIR lissés de cancers du rein chez les hommes selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

L'analyse de sensibilité prenant en compte les informations disponibles sur les cas de cancer du rein pris en charge par les urologues ayant refusé de participer au recueil de données ne montre aucune variation au-delà des variations aléatoires (Annexe 11).

### *Identification de clusters*

Les méthodes globales de clustering montrent une hétérogénéité spatiale liée à une surdispersion des SIR de cancer du rein (Potthoff-Whittinghill ;  $p = 0,0018$ ) mais pas de corrélation spatiale globale ni de regroupements de cas à l'intérieur de la zone d'étude.

La recherche de clusters par statistique de scan n'a pas permis de détecter de clusters significatifs.

Ces résultats témoignent d'une répartition globalement homogène des cas de cancers du rein sur le territoire.

### **B) Résultats chez les femmes**

Vingt-trois communes, représentant 2 % de la population féminine du département, ne comptent aucun cas de cancer du rein diagnostiqué chez les femmes entre 2013 et 2018.

Trois cent soixante-six Iris ne recensent aucun cas de cancer du rein chez la femme sur la période 2013-2018 (45 %), ces Iris représentent 38 % de la population totale féminine du département. Comme chez l'homme, les analyses spatiales n'ont pas été réalisées à cette échelle.

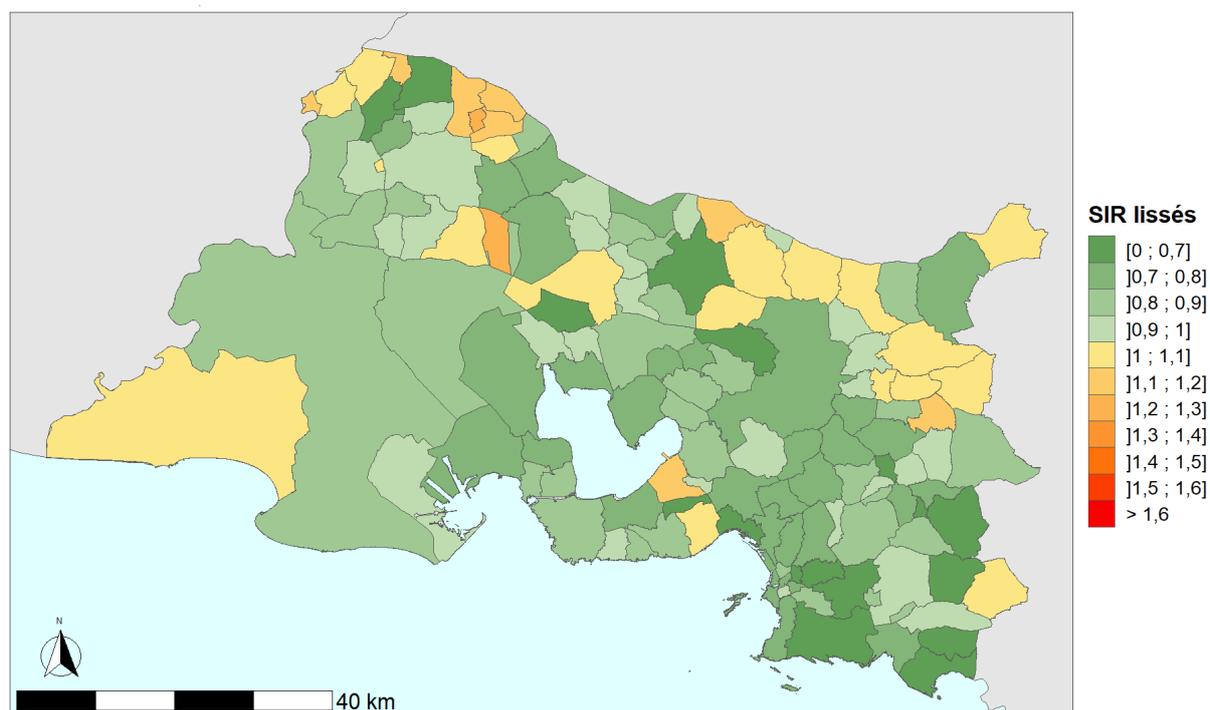
### *Cartographie des risques lissés*

Les résultats du modèle de lissage BYM ajusté sur l'ensemble des covariables montrent un effet négatif significatif du tabagisme ( $RR = 0,23$  IC95 % [0,08 ; 0,73]), ainsi qu'un effet négatif non significatif de la densité de population sur la surincidence de cancer du rein ( $RR = 0,79$  IC95 % [0,59 ; 1,08] et  $RR = 0,76$  IC95 % [0,57 ; 1,06] respectivement pour les niveaux de l'indice de densité de l'Insee de densité intermédiaire et de forte densité par rapport au niveau de faible densité) (Annexe 10).

L'absence d'autocorrélation des résidus du modèle sélectionné est confirmée par un indice de Moran faible ( $p = 0,99$ ).

Les SIR lissés varient de 0,61 à 1,26 en médiane *a posteriori* avec une médiane de 0,86. Ils sont globalement inférieurs à 1 sur le département. Les valeurs supérieures à 1 se concentraient dans la moitié nord (Figure 17). Aucune commune ne présente de surincidence significative. Néanmoins, un regroupement de 4 communes au nord, autour de Verquières, présente des surincidences supérieures à 1,2 à la limite de la significativité. Au contraire, 28 communes présentent une sous-incidence significative par rapport au niveau national, dont onze arrondissements de la ville de Marseille.

**Figure 17. Cartographie des SIR lissés de cancers du rein chez les femmes selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

L'analyse de sensibilité prenant en compte les informations disponibles sur les cas de cancer du rein pris en charge par les urologues ayant refusé de participer au recueil de données ne montre aucune variation au-delà des variations aléatoires (Annexe 11).

#### *Identification de clusters*

Les méthodes globales de clustering montrent une hétérogénéité spatiale liée à une surdispersion des SIR de cancer du rein (Potthoff-Whittinghill ;  $p < 10^{-4}$ ) mais pas de corrélation spatiale globale ni de regroupements de cas à l'intérieur de la zone d'étude

La recherche de clusters par statistique de scan n'a pas permis la détection de clusters significatifs.

Ces résultats témoignent d'une répartition globalement homogène sur le territoire des cas de cancers du rein chez les femmes.

### 4.2.3 Leucémies aiguës myéloïdes

#### 4.2.3.1 Incidence départementale estimée

L'incidence départementale des LAM standardisée monde tous âges confondus dans le département des Bouches-du-Rhône est estimée à 3,0 cas pour 100 000 personnes-années chez les hommes et à 2,5 cas pour 100 000 personnes-années chez les femmes, soit des taux d'incidence similaires à ceux observés au niveau de la France (Tableau 12).

**Tableau 12. Taux d'incidence bruts et standardisés monde (TSM) pour les LAM pour 100 000 personnes-années, Bouches-du-Rhône 2013-2018 (incluant les cas de moins de 20 ans)**

Sexe	Bouches-du-Rhône 2013-2018			France 2013-2018	
	Nombre de cas	Taux brut	TSM [IC 95 %]	Taux brut	TSM [IC 95 %]
Femmes	305	4,8	2,5 [2,1 ; 2,9]	4,6	2,3 [2,2 ; 2,3]
Hommes	311	5,4	3,0 [2,6 ; 3,5]	5,4	3,0 [3,0 ; 3,1]

#### 4.2.3.2 Répartition spatiale des risques

##### A) Résultats chez les hommes

Cinquante-deux communes, représentant 10 % de la population totale masculine du département, ne comptent aucun cas de LAM sur la période 2013-2018.

Cinq cent soixante-trois Iris (70 %) ne comptabilisent aucun cas de LAM chez l'homme sur la période 2013-2018, ces Iris représentaient 65 % de la population totale masculine du département. Compte tenu du grand nombre d'Iris sans aucun cas de LAM recensé, les analyses spatiales n'ont pas été réalisées à cette échelle.

##### Cartographie des risques lissés

Les résultats du modèle de lissage ajusté sur l'ensemble des covariables ne montrent aucun effet significatif des covariables sur les valeurs des SIR (Annexe 10).

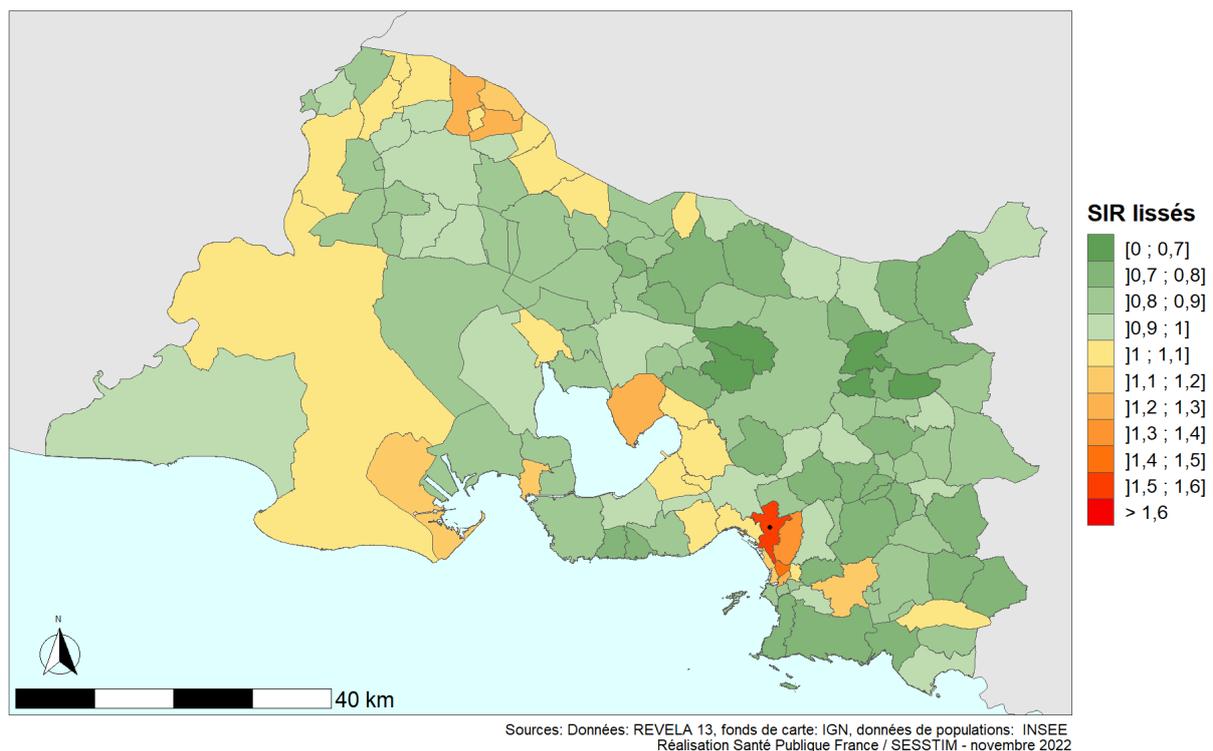
L'absence d'autocorrélation des résidus est confirmée par un indice de Moran faible ( $p = 0,83$ ).

L'interaction entre le *proxy* du tabagisme et l'indice Fdep étant non significative, elle n'a pas été retenue dans le modèle.

Les SIR lissés varient de 0,56 à 1,52 en médiane *a posteriori* avec une médiane de 0,88. Ils sont globalement inférieurs à 1 sur le département. Cependant, un premier regroupement a été identifié au nord de Marseille avec une surincidence significative dans le 15<sup>e</sup> arrondissement et une surincidence élevée dans les 3<sup>e</sup> et 14<sup>e</sup> arrondissements (Figure 18). Un autre regroupement incluant les communes rurales de Noves, Cabannes, Verquières et Saint-Andiol est également observé au nord-ouest du département à la limite du département du Vaucluse.

Ces résultats sont à interpréter avec précaution en raison du faible nombre de cas et de la largeur des intervalles de crédibilité

**Figure 18. Cartographie des SIR lissés de LAM chez les hommes selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



### *Identification de clusters*

À l'échelle des communes, la répartition spatiale des LAM est homogène sur le département tant en termes de surdispersion, comme l'atteste le test de Potthoff-Whittinghill ( $p = 0,56$ ), qu'en termes de corrélation spatiale, avec un indice de Moran ( $p = 0,62$ ), un EBI ( $p = 0,81$ ) et un test de Tango ( $p = 0,76$ ) non significatifs.

La recherche de clusters par statistique de scan n'a pas permis de détecter de clusters significatifs à une taille maximale des clusters autorisée fixée à 30 %.

### **B) Résultats chez les femmes**

Cinquante-huit communes ne comptent aucun cas de LAM sur la période 2013-2018, représentant 12 % de la population totale féminine du département.

Cinq cent soixante-seize Iris (71 %) ne recensent aucun cas de LAM chez les femmes sur la période 2013-2018, ces Iris représentent 67 % de la population totale féminine du département. Compte tenu du grand nombre d'Iris sans aucun cas de LAM recensé, les analyses spatiales n'ont pas été réalisées à cette échelle.

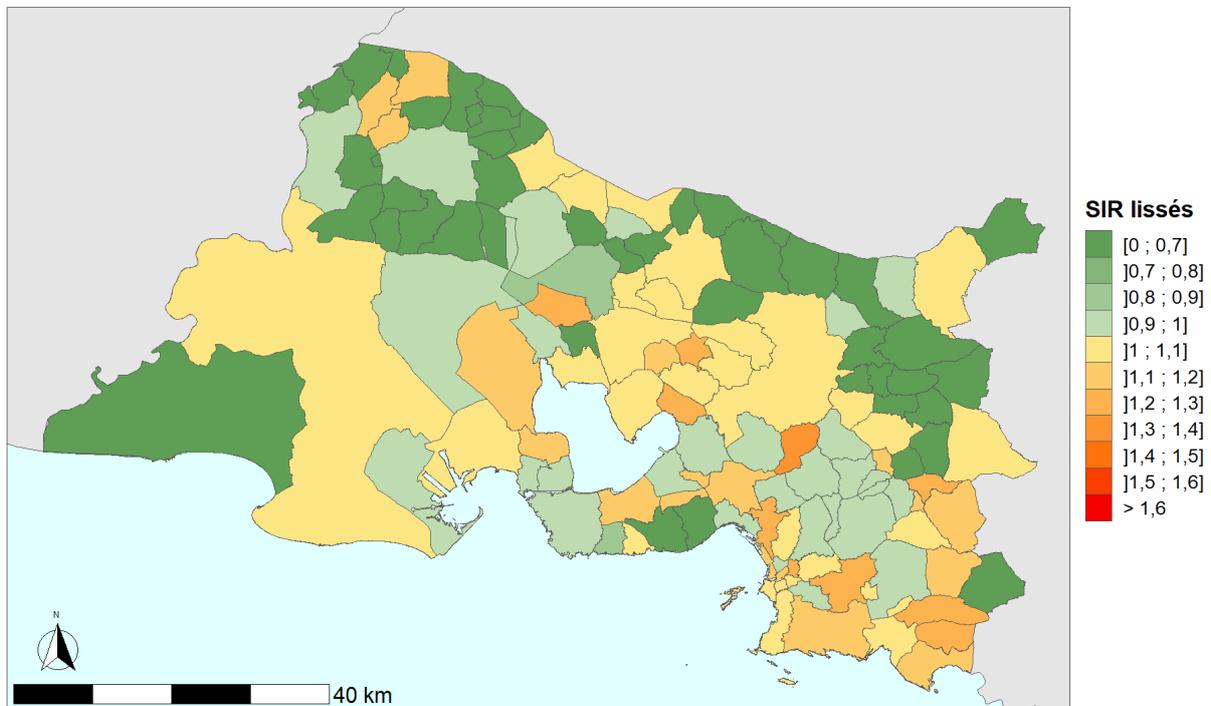
### *Cartographie des risques lissés*

Les résultats du modèle de lissage ajusté sur l'ensemble des covariables montrent un impact significatif du niveau de densité de population pour les communes de densité intermédiaire (RR = 2,21 IC95 % [1,22 ; 4,78]) et de densité élevée (RR = 2,10 IC95 % [1,16 ; 4,70]) par rapport aux communes de densité faible (Annexe 10).

L'absence d'autocorrélation des résidus est confirmée par un indice de Moran faible ( $p = 0,87$ ).

Les SIR lissés varient de 0,39 à 1,35 en médiane *a posteriori* avec une médiane de 0,97. Un nombre important de communes présente des SIR inférieurs à 0,7 sur la moitié nord du département et une part importante de communes présente des valeurs proches de 1. Le quart sud-est concentre la plupart des communes ayant un SIR estimé au-delà de 1,1 (Figure 19). Aucune surincidence significative n'est observée.

**Figure 19. Cartographie des SIR lissés de LAM chez les femmes selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

### Identification de clusters

À l'échelle des communes, la répartition spatiale des LAM est homogène sur le département tant en termes de surdispersion, comme l'atteste le test de Potthoff-Whittinghill ( $p = 0,87$ ), qu'en termes de corrélation spatiale, avec un indice de Moran ( $p = 0,76$ ), un EBI ( $p = 0,74$ ) et un test de Tango ( $p = 0,46$ ) non significatifs.

La recherche de clusters par statistique de scan n'a pas permis de détecter de clusters significatifs à une taille maximale des clusters autorisée fixée à 30 %.

L'analyse de sensibilité de l'ensemble des cas de LAM hommes et femmes confondus est présentée en Annexe 12.

## 5 DISCUSSION

### 5.1 Synthèse des résultats

Au niveau départemental, l'incidence du cancer de la vessie sur la période 2013-2018, pour les hommes comme pour les femmes, est significativement supérieure à celle estimée au niveau national sur la même période, avec des taux standardisés monde pour 100 000 personnes-années respectivement de 18,8 et 3,0 versus 14,4 et 2,3 en France métropolitaine.

Après lissage des incidences standardisées, la cartographie des SIR montre des incidences significativement plus élevées que celles attendues dans 31 communes du département pour les hommes (23 %), dont les 16 arrondissements de Marseille, et dans 13 arrondissements de Marseille pour les femmes, soit 10 % des communes. De plus, la recherche de clusters par la statistique de scan spatial a mis en évidence des agrégats de commune avec un risque relatif plus élevé sur ces mêmes zones que pour le reste des communes, pour les deux sexes. Une recherche de clusters à l'échelle de l'Iris, conduite uniquement chez les hommes, n'a pas permis de détecter d'autre regroupement de cas sur le département, en dehors des clusters identifiés à la commune mais recentrés sur quelques Iris de ces mêmes communes, ce qui conforte les résultats obtenus à l'échelle de la commune.

Les incidences départementales estimées du cancer du rein et des LAM ne sont pas différentes de l'incidence nationale pour les deux sexes, sur la période 2013-2018. Les niveaux d'incidence de cancer du rein chez les hommes comme chez les femmes sont répartis de façon hétérogène sur le territoire mais sans regroupement de cas identifiable, que ce soit par les méthodes globale ou locale de clustering. Concernant les LAM, les cas sont répartis de façon globalement homogène sur le territoire. Cependant, un arrondissement de Marseille présente une surincidence significative, uniquement chez les hommes. Aucun cluster n'a été détecté sur le département tant chez les hommes que chez les femmes.

### 5.2 Cancers de la vessie

#### 5.2.1 Caractéristiques sociodémographiques et histologiques

Les caractéristiques sociodémographiques des cas de cancers de la vessie inclus dans l'Observatoire REVELA13 sont similaires à celles observées en France métropolitaine avec un âge médian au diagnostic légèrement plus élevé pour les deux sexes, 75 ans pour les hommes et 79 ans pour les femmes versus respectivement 73 et 78 ans au niveau national et un sexe ratio (H/F) à 4,6 versus 4,4 en France métropolitaine (20). Ces caractéristiques sont similaires à celles de la première étude publiée en 2019 (6).

Comme en 2019, les femmes sont diagnostiquées à un âge plus tardif que les hommes et présentent au diagnostic des tumeurs à un stade plus avancé (fréquence plus élevée de TVIM). Ces constats, également retrouvés dans la littérature, ne s'expliqueraient pas par des différences d'exposition aux facteurs de risque mais par des mécanismes moléculaires différents, notamment au niveau du métabolisme des agents cancérigènes par les enzymes hépatiques, ainsi que par le rôle différentiel joué par les hormones stéroïdiennes sexuelles (49). Les récepteurs aux œstrogènes joueraient ainsi un rôle dans le développement du cancer de la vessie entraînant des effets biologiques différents en fonction du sexe (50). Enfin, le diagnostic des femmes à un stade plus avancé serait également lié au délai de prise en charge de celles-ci du fait d'une différence de traitement et de protocole de prise en charge des hématuries (51). Une sensibilisation des médecins traitants, des gynécologues et des

urologues mais aussi des patientes à ce sujet semble aujourd'hui indispensable pour améliorer la prise en charge de ce cancer chez la femme (52,53).

## 5.2.2 Cartographie des risques et détection de clusters :

Comme dans le premier rapport (6), la répartition des cancers de la vessie est hétérogène sur le département avec une tendance aux regroupements de cas chez les hommes et chez les femmes. En effet, l'augmentation du nombre de cas enregistrés a permis d'augmenter la puissance des analyses et cette tendance est également observée chez les femmes.

Chez les hommes, après prise en compte de ces facteurs d'ajustement, la zone de surincidence observée précédemment à l'est du département, englobant la ville de Marseille et les communes voisines dans sa partie sud-est et nord, est toujours présente. Cependant, cette nouvelle étude a également mis en évidence un surrisque significatif de cancer de la vessie pour quatre communes situées sur le pourtour de l'étang de Berre (Port-de-Bouc, Marignane, Vitrolles et Sausset-les-Pins).

La cartographie des risques lissés à l'échelle des Iris confirme la surincidence significative observée dans l'ensemble de la ville de Marseille (hormis dans le nord du 15<sup>e</sup> arrondissement) et dans le sud-est du département mais pas sur le pourtour de l'étang de Berre.

Les clusters identifiés sont similaires à l'étude précédente. Le cluster principal englobe les communes du sud-est du département et la partie est de Marseille (RR = 1,37) et le deuxième concerne tous les autres arrondissements de Marseille à l'exception du 15<sup>e</sup> arrondissement (RR = 1,31). La recherche de clusters à l'échelle de l'Iris restreint les clusters observés à l'échelle communale à quelques Iris de ces communes mais ne permet pas de détecter d'autres clusters sur le département.

Chez les femmes, contrairement aux résultats de 2019, une surincidence significative par rapport au niveau national est observée pour l'ensemble des arrondissements de la ville de Marseille, à l'exclusion des 3<sup>e</sup>, 14<sup>e</sup> et 16<sup>e</sup> arrondissements. La recherche de clusters identifie également un agrégat de communes présentant un excès de risque de cancer de la vessie sur une zone à peu près identique (RR = 1,48).

La répartition géographique hétérogène des cancers de la vessie et la tendance à l'agrégation sont bien décrites chez les hommes dans la littérature et ont fait l'objet d'une discussion détaillée dans le premier rapport de l'Observatoire REVELA 13. Les études chez les femmes sont plus rares car les nombres de cas de cancers de la vessie étudiés sont toujours plus faibles. Une étude récente de Saint-Jacques *et al* retrouve également une hétérogénéité spatiale assez similaire chez les femmes et les hommes mais avec une surincidence plus marquée chez ces derniers (54).

En France, en 2015, 39 % des cas de cancers de la vessie étaient attribuables au tabagisme chez les hommes et 13 % chez les femmes. Le cancer de la vessie était ainsi, après le cancer du poumon, du larynx et de la cavité orale, l'un des cancers qui contribuait le plus au nombre de cas de cancers attribuables au tabagisme actif (55). Le lien entre tabagisme et risque de cancer de vessie est vérifié dans notre étude avec une association significativement positive retrouvée à l'échelle de la commune chez les hommes, qui persiste sans être significative à l'échelle de l'Iris. Chez les femmes, une association est également retrouvée (RR = 2,13), même si celle-ci n'est pas significative, du fait sans doute d'effectifs trop faibles. Cependant, la prise en compte de ce facteur de confusion ne suffit pas à expliquer les zones de surincidences retrouvées sur le département.

D'autres facteurs étudiés dans la littérature semblent jouer un rôle dans la survenue des cancers de vessie et/ou dans leur répartition géographique à l'image de la défaveur sociale et de la densité de population (9,56). Dans notre étude, aucune association positive n'est retrouvée avec l'indice de défaveur sociale retenu (Fdep), tant chez les hommes que chez les femmes, et quelle que soit l'échelle géographique considérée. Ce résultat peut s'expliquer par l'influence plus importante de l'autocorrélation spatiale dans la répartition des cas de vessie, comme l'a montré l'équipe de Goungounga *et al* en 2016, contrairement à ce qu'ils avaient observé pour le cancer du poumon, dont la répartition spatiale était très liée à la défaveur sociale (56). Par ailleurs, la défaveur sociale mesurée par le Fdep est relativement plus importante dans le département, comparativement au niveau national, avec plus de 40 % des communes dans le quintile le plus défavorisé et plus de 80 % dans les trois premiers. La relative homogénéité de la défaveur sociale sur le département pourrait engendrer un défaut de contraste entre les communes, ce qui pourrait expliquer en partie l'absence d'association apparente. Le recours à la mortalité par cancer du poumon pour modéliser le tabagisme, couplé à sa corrélation avec la défaveur sociale, peut également expliquer l'absence d'effet observé de la défaveur sociale, dissimulé par celui du tabagisme. Pour autant, l'intégration de l'interaction entre ces deux facteurs au modèle de lissage ne modifie pas les résultats.

La mise à jour des analyses spatiales présentée dans ce rapport met en évidence une association positive entre la densité de population et le cancer de la vessie pour les deux sexes, même si elle n'est significative que chez les hommes. Le risque de cancer de vessie est en effet plus important dans les communes plus denses comparées aux communes de plus faible densité. Cette association n'a pas été observée lors des premières analyses spatiales publiées en 2019. Cependant, ce facteur n'est pas pris en compte de la même façon dans les deux études. En 2019, c'est la densité brute de population qui était intégrée aux analyses spatiales, alors que dans cette seconde étude, la densité a été remplacée par la grille de densité communale définie par l'Insee, reflétant davantage les degrés d'urbanisation.

L'influence de la densité de population sur le risque de cancer a été étudiée en Amérique du Nord dès les années 80. Ces études ont montré que l'incidence de nombreux cancers est liée de façon linéaire à la densité de population et notamment pour le cancer de la vessie (57,58). Plus récemment, Radespiel *et al* ont observé un risque accru de cancers, dont le cancer de la vessie pour les deux sexes, dans les centres-villes, comparés aux zones suburbaines ou rurales (59). Castano *et al*, dans une enquête cas-témoins portant sur plus de 24 000 sujets, ont montré que le fait de résider dans une ville de plus de 100 000 habitants pendant quarante ans était associé à un risque accru significatif de cancer de vessie (OR=1,3) (60). Dans notre étude, les surincidences les plus importantes sont également retrouvées sur Marseille, principale agglomération du département.

Si la densité n'est pas un facteur de risque de cancer, elle est cependant le reflet d'expositions communes qui lui sont associées, tant au niveau des modes de vie qu'au niveau des expositions environnementales, comme la pollution atmosphérique. En 2013, les experts du CIRC ont classé l'exposition à la pollution atmosphérique dans le groupe 1 des cancérogènes avérés pour l'homme, sur la base de preuves suffisantes concernant un lien avec le risque accru de cancer du poumon. Ils ont également noté une association positive avec un risque accru de cancer de la vessie<sup>9</sup>. Ces dix dernières années, de nombreuses études ont été menées pour mettre en évidence un lien entre pollution atmosphérique et incidence et/ou mortalité par cancer de vessie, au travers de différents indicateurs de pollution, particulièrement les particules et les oxydes d'azote. Elles ne sont pas concordantes et une récente revue de la littérature a même conclu sur un niveau de risque faible voir très faible (61).

---

<sup>9</sup> [https://www.iarc.who.int/wp-content/uploads/2018/07/pr221\\_F.pdf](https://www.iarc.who.int/wp-content/uploads/2018/07/pr221_F.pdf)

Le département des Bouches-du-Rhône est un territoire soumis à d'importantes émissions de polluants atmosphériques notamment du fait de la présence de grandes zones urbanisées, de réseaux routiers et autoroutiers denses, du pôle industriel de Fos-sur-Mer et du port de Marseille. Les polluants liés au trafic automobile sont particulièrement présents dans les zones urbaines denses et pourraient expliquer en partie la concentration de cancer de la vessie qui y est observée.

Enfin, l'exposition professionnelle constitue l'autre des facteurs de risque du cancer de la vessie le plus important avec le tabagisme (62). Le département des Bouches-du-Rhône<sup>10</sup> concentrait, en 2018, 35,6 % des industries et 50 % des emplois industriels de la région, ce qui en faisait le 1<sup>er</sup> département industriel de la région PACA. L'Observatoire REVELA13 ne recueille pas de données individuelles sur les professions des patients qui ne sont pas systématiquement renseignées dans les dossiers médicaux ou, lorsqu'elles existent, ne sont pas suffisamment précises pour estimer des éventuelles expositions à des cancérigènes professionnels. L'influence des expositions professionnelles n'est donc pas prise en compte dans nos analyses.

## 5.3 Cancers du rein

### 5.3.1 Caractéristiques sociodémographiques et histologiques

Les caractéristiques d'âge et de sexe pour les cancers du rein inclus dans l'Observatoire sont globalement similaires à celles de la littérature et plus particulièrement avec les dernières données publiées par Santé publique France (20) avec un sexe ratio H/F à 2,2 et un âge médian au diagnostic légèrement inférieur tant chez les hommes que chez les femmes, respectivement 65 et 67 ans versus 67 et 70 ans au niveau national. Aucune différence n'est observée entre les hommes et les femmes concernant le stade et le grade des tumeurs. Concernant les formes histologiques des tumeurs, les hommes présentent plus fréquemment des carcinomes tubulo-papillaires, alors qu'à l'opposé les carcinomes chromophobes sont plus fréquents chez les femmes. Cette différence est connue des urologues et précisé dans le descriptif des différentes formes histologiques des cancers à cellules rénales<sup>11</sup>.

### 5.3.2 Cartographie des risques et détection de clusters

Aucune commune ne présente de surincidence significative pour le cancer du rein, même si quelques SIR compris entre 1 et 1,3 ont été retrouvés, notamment chez les hommes, dans le 16<sup>e</sup> arrondissement de Marseille et, chez les femmes, dans la partie située au nord-ouest du département.

À l'échelle du département, la répartition géographique des cancers du rein chez les hommes et les femmes est hétérogène mais sans tendance au regroupement. Aucun cluster n'a été détecté à l'échelle de la commune. La répartition géographique du cancer du rein a été moins étudiée dans la littérature que pour le cancer de la vessie. Une hétérogénéité spatiale a parfois été retrouvée chez les hommes (54,63). L'absence de cluster de cancer du rein est également rapportée dans certaines études (64,65) alors que d'autres identifient des clusters plutôt urbains (16,66).

Les facteurs de risque avérés du cancer du rein sont le tabagisme, l'obésité, l'hypertension artérielle, un antécédent de maladie rénale chronique ou de maladie génétique, l'exposition

<sup>10</sup> [http://www.orspaca.org/sites/default/files/portrait\\_bouches-du-rhone\\_orspaca.pdf](http://www.orspaca.org/sites/default/files/portrait_bouches-du-rhone_orspaca.pdf)

<sup>11</sup> <https://www.arcagy.org/infocancer/localisations/rein-et-voies-urinaires/cancer-rein/formes-de-la-maladie/les-cancers-affectant-le-rein.html/>

professionnelle au trichloréthylène (67). Ce sont essentiellement des facteurs individuels non recueillis dans le cadre de l'Observatoire REVELA13. Les études s'intéressant aux expositions à des facteurs environnementaux sont encore peu conclusives.

Chez les hommes, des effets positifs du tabagisme et de la densité de population sont observés dans notre étude, sans que ceux-ci ne soient significatifs, alors que chez les femmes, un effet négatif significatif avec le tabagisme est observé. Ce résultat est surprenant du fait du lien avéré entre tabagisme actif et risque de cancer du rein (68), même si les risques relatifs associés au statut tabagique sont inférieurs à ceux observés pour le cancer de la vessie (69). En France, en 2015, 38 % des cas de cancers du rein étaient en effet attribuables au tabagisme chez les hommes et 8 % chez les femmes (55). En revanche, avoir arrêté de fumer depuis plusieurs années diminuerait le risque (67). L'utilisation d'un *proxy* pour estimer le tabagisme et les évolutions de profil de tabagisme chez les femmes pourrait être des pistes d'explications de ce résultat.

Dans notre étude, les incidences les plus élevées du cancer du rein chez la femme sont plutôt observées sur les communes du nord-ouest du département, alors que les arrondissements de Marseille et les communes du sud-est du département, où le tabagisme est également important, présentent plutôt un profil de sous-incidence. Ces résultats sont sans doute à l'origine de l'effet négatif significatif du tabagisme observé chez les femmes.

Aucune association n'a été mise en évidence avec la densité de population, l'accès aux soins ou la défaveur sociale. Cette absence d'association avec le statut socio-économique que ce soit dans sa dimension matérielle ou sociale est également retrouvée dans la littérature (54,70).

## 5.4 Leucémies aiguës

### 5.4.1 Caractéristiques sociodémographiques et histologiques

Les LAL surviennent surtout chez l'enfant entre 2 à 15 ans (où elles constituent le plus fréquent des cancers) et sont plus rares chez l'adulte, en particulier au-delà de 50 ans. À l'inverse, les LAM surviennent chez le petit enfant de moins de 2 ans, puis au-delà de 15 ans, où elles ont une fréquence progressivement croissante. L'âge médian des LAM se situe autour de 70 ans, avec des cas de plus en plus nombreux au-dessus de cet âge compte tenu du vieillissement progressif de la population. Au niveau national, l'âge médian au diagnostic est de 69 ans chez l'homme et de 72 ans chez la femme, il est de 71 ans pour les deux sexes dans notre étude. Le sexe ratio H/F proche de 1 est inférieur à celui observé au niveau national (1,35) (21) mais similaire à celui du registre des hémopathies de Gironde<sup>12</sup>.

Depuis 1990, l'augmentation de l'incidence est plus marquée chez les femmes les plus âgées, augmentation qui semble indépendante de l'évolution des facteurs démographiques. En effet, la part attribuable aux facteurs de risque s'élève à 69 % chez la femme. Dans notre étude, la répartition des cas en fonction du sexe et des classes d'âge fait nettement apparaître un accroissement du nombre de cas chez les femmes à partir de 85 ans. En l'absence d'études spécifiques à grande échelle, les facteurs pouvant expliquer cette augmentation chez la femme âgée n'ont pas été identifiés (21).

L'évolution des classifications internationales permet de distinguer actuellement quatre grandes catégories de LAM : les formes avec anomalies cytogénétiques récurrentes, les formes avec myélodysplasie, celles secondaires à une thérapeutique cytotoxique et les autres

<sup>12</sup> [https://sites.bph.u-bordeaux.fr/REGISTRES-CANCERS-AQUITAINE/HM/H\\_Resultats.aspx](https://sites.bph.u-bordeaux.fr/REGISTRES-CANCERS-AQUITAINE/HM/H_Resultats.aspx)

formes. Que ce soit en France ou dans les autres pays, la plupart des études épidémiologiques concernent encore l'ensemble des LAM (71-73) ou le sous-groupe des LAM de novo (exclusion des LAM secondaires) (74). Elles incluent généralement les enfants et du fait de méthodes d'enregistrements des LAM différentes selon les pays, la comparaison avec nos résultats reste difficile. La codification selon les classifications OMS 2008 et 2016 étant récente, il faudra encore plusieurs années avant de pouvoir étudier un nombre de cas suffisant correspondant aux différentes entités de ces classifications tant en France qu'à l'étranger.

#### 5.4.2 Cartographie des risques et détection de clusters

La cartographie des SIR lissés chez les hommes a mis en évidence un arrondissement de Marseille avec une surincidence significative de LAM (15<sup>e</sup> arrondissement). Aucune surincidence significative n'a été observée chez les femmes. De plus, l'analyse de sensibilité réalisée en regroupant les hommes et les femmes dans le même modèle n'a pas permis d'identifier d'autres communes en surincidence.

Les cas de LAM sont répartis de façon globalement homogène sur le territoire et ne présentent pas de tendance au regroupement de cas. Aucun cluster n'a été détecté à la commune pour les deux sexes et dans l'analyse conjointe.

Les études portant sur la répartition géographique des LAM ou leucémies aiguës sont pratiquement toutes réalisées chez les enfants de moins de 15 ans comme cela a été rappelé dans le premier rapport de l'Observatoire publié en 2019. À l'échelle de la France, l'incidence tous âges des LAM n'est connue que pour les départements couverts par un registre de cancer. La représentation cartographique des SIR observés dans ces départements montre qu'il existe bien des différences géographiques d'incidence en France métropolitaine, même si celles-ci ne sont pas significatives (31). Enfin, les rares études portant sur l'analyse des leucémies aiguës n'ont pas mis en évidence de cluster de leucémies myéloïdes ou lymphoïdes (65).

Un effet de la densité de population a été identifié chez les femmes, le risque de développer une LAM est en effet plus important dans les communes densément peuplées (RR = 2,10) ou de densité intermédiaire (RR = 2,21), comparé aux communes de faible densité. Ces effets importants s'expliqueraient en partie par un biais dû au très faible taux d'incidences de la maladie sur la période considérée, rendant l'apparition de cas dans des zones peu peuplées proches de zéro. Ces zones se confondent en grande partie avec les zones de niveau de densité faible. Ce lien avec la densité de population est peu étudié. Le registre des hémopathies de Côte d'Or rapporte que le nombre de cas est deux fois plus élevé en zone urbaine (75) et Sasaki *et al* retrouvent un nombre plus élevé de cas dans les agglomérations (74).

Un effet du tabagisme a également été observé, et ce uniquement chez les femmes, mais il est non significatif et présente un très large intervalle de crédibilité.

Aucune association n'a été mise en évidence avec l'accès aux soins ou la défaveur sociale. Une étude danoise a analysé le lien entre niveau socioéconomique et LAM mais uniquement chez les parents d'enfants malades (76).

L'origine de la maladie reste encore largement méconnue en dehors de quelques facteurs de risque bien identifiés comme le tabagisme, l'exposition à des agents physiques ou chimiques (radiations ionisantes, benzène), la présence de maladies génétiques (syndrome de Down, anémie de Fanconi...), d'antécédents de cancers solides et d'hémopathies malignes, ou encore le fait d'avoir reçu un traitement par chimiothérapie ou par radiothérapie (21).

## 5.5 Discussion des choix méthodologiques

### 5.5.1 Choix des variables d'ajustement

#### 5.5.1.1 *Défaveur sociale*

Le choix de l'utilisation et de la modélisation du Fdep comme indice de défaveur sociale avait été discuté dans la précédente étude de 2019, après comparaison avec deux autres indices largement utilisés dans la littérature, l'EDI et l'indice de Townsend. Cependant, au vu de ces nouveaux résultats, et en l'absence de toute association avec les cancers étudiés, le choix de cet indice pourrait être revu du fait de son caractère peu discriminant sur le département (cf 5.2.2). D'autres indices permettant de caractériser la défaveur sociale pourraient être testés dans de futures analyses spatiales, tels ceux issus des travaux menés par l'équipe du Sesstim dans le cadre de l'épidémie de Covid-19 pour caractériser localement les profils socio-économiques des Iris de la région PACA (77), bien que l'utilisation de profils construits de manière *ad hoc* rende difficile la comparaison de résultats d'une étude à l'autre.

#### 5.5.1.2 *Densité de population*

Bien qu'il soit difficile d'imaginer un lien causal direct entre l'apparition d'un cancer à l'échelle individuelle et la densité de population du territoire de résidence, cette dernière est largement utilisée dans les approches populationnelles pour représenter des facteurs de risque socio-économiques, comportementaux et environnementaux, liés notamment aux statuts rural ou urbain des territoires (78,79). Cependant, la modélisation de la densité se heurte à la problématique du choix de la relation *a priori*.

L'utilisation d'une variable continue et d'une relation linéaire, comme cela avait été fait lors des premières analyses publiées en 2019, n'avait pas permis de discriminer correctement les territoires péri-urbains des territoires fortement urbanisés, et à l'inverse, les territoires peu denses des territoires ruraux. Une catégorisation en quartiles a donc été réalisée plutôt qu'une modélisation logarithmique qui n'aurait pas permis autant de flexibilité (en faisant l'hypothèse d'une augmentation du risque de cancer pour des densités de population très élevées souvent associées à des catégories de populations défavorisées). Cette modélisation a été utilisée à l'échelle des Iris.

Une autre problématique est apparue à l'échelle communale avec une importante variabilité du tissu urbain, et donc de la densité, à l'intérieur des communes. L'indice de densité de population de l'Insee semble apporter une réponse utile à cette problématique en considérant non pas la densité moyenne mais la répartition réelle des habitants dans des zones plus ou moins densément peuplées, transcrivant ainsi une meilleure estimation de l'environnement de la population (7).

Cet indice de densité n'est pas disponible à l'échelle des Iris. Toutefois, la définition même des Iris est basée sur une homogénéité du territoire concerné et possède une distribution de la population plus régulière. Néanmoins, le découpage en quartiles pourrait être retravaillé pour créer des classes d'Iris plus proche des classes de communes de l'indice de densité de l'Insee, en se référant aux seuils utilisés pour sa construction (7).

#### 5.5.1.3 *Accès aux soins de premier recours*

L'intégration de ce facteur dans les analyses reposait sur l'hypothèse que l'accès aux soins peut influencer la précocité du diagnostic et prévenir ainsi l'apparition d'une tumeur maligne ou sa prise en charge à un stade précoce. Néanmoins, la littérature fournit peu d'études probantes sur le lien entre l'accès aux soins et la morbi-mortalité des cancers (80-82).

Le choix de l'APL pour modéliser l'accès aux soins a l'avantage d'intégrer notamment une meilleure définition de l'offre et de la demande de soins, en prenant en compte l'activité des professionnels de santé et la structure d'âge de la population pouvant recourir à des soins. Le calcul de l'APL à l'échelle communale tient compte à la fois de l'offre et de la demande issues de la commune et des communes voisines ainsi que du recours différencié aux soins, en affectant un poids plus important aux tranches d'âge les plus consommatrices de soins. Cet indice est apparu plus pertinent pour refléter l'accès réel aux soins de premiers recours que l'offre de soins, illustrée par la seule densité de médecin, en caractérisant plus finement les situations auxquelles font face les habitants. Cependant, si l'utilisation de l'APL reflète mieux l'accès réel aux soins, il ne renseigne pas sur le recours réel aux soins, fortement lié aux inégalités socio-économiques des individus.

Pour autant, les résultats des modèles de lissage pour les trois cancers étudiés et les deux sexes n'ont pas montré d'associations entre l'APL et l'incidence de ces cancers.

Sans données disponibles à l'Iris, les valeurs communales ont été utilisées. Ce facteur incluant une dimension spatio-temporelle pour définir l'accessibilité, cette estimation entraînait inévitablement un biais, en particulier pour les communes les plus grandes. À défaut de meilleurs indicateurs disponibles à cette échelle, la densité de médecin généralistes pourrait être testée (82).

Lors de la première étude conduite en 2019, une association positive avait été retrouvée chez les femmes entre le risque de cancer de la vessie et l'accès à un urologue, calculé à partir du nombre d'urologues exerçant à une distance de moins de 30 minutes du centroïde de la commune de résidence. Cet indicateur avait été construit spécifiquement pour l'étude sans validation externe et n'avait pas pu être utilisé pour les analyses des leucémies aiguës. Lors de cette seconde étude, la mise à jour de cet indicateur s'est avérée complexe du fait de l'évolution constante de l'offre en urologie sur le territoire. La complexité de sa mise à jour et les imprécisions inhérentes à cet indice nous ont conduits à l'abandonner au profit d'un indice validé d'accès aux soins, qui pouvait être introduit dans les analyses spatiales pour les trois cancers étudiés. Plus que l'accès au spécialiste, c'est l'accès à une densité de consultations spécialisées en urologie qui serait intéressante mais cette donnée reste difficile à estimer.

#### 5.5.1.4 Tabagisme

Le tabagisme est un facteur de risque de nombreux cancers et reste le principal facteur de risque du cancer de la vessie. Or, la région PACA présente une proportion très élevée de fumeurs. L'introduction de ce facteur de risque dans les analyses avait été proposée lors de la précédente étude. Cependant, la prévalence du tabagisme n'est pas disponible à l'échelle de la commune et le statut tabagique des cas n'est pas recueilli par l'Observatoire, n'étant pas systématiquement disponible dans les dossiers médicaux. Un *proxy* du tabagisme a donc été recherché dans la littérature (15-17).

L'utilisation de la mortalité par cancer du poumon comme *proxy* du tabagisme se justifiait par la relation causale forte et bien établie entre les deux. Elle offre en outre l'avantage majeur de se baser sur des données fiables et disponibles. Les données désagrégées par sexe ont permis de révéler une structure différente de la répartition spatiale du *proxy* chez les hommes et les femmes. En outre, l'usage de la mortalité par cancer implique une exposition au tabac bien antérieure à la date de décès, lié au délai d'apparition de cancers du poumon. L'index de tabagisme ainsi estimé correspond à une période antérieure à notre période d'étude, mais compatible avec d'autres cancers ayant un délai d'apparition similaire, comme le cancer de la vessie.

L'indisponibilité des données de décès par cancer du poumon à l'échelle de l'Iris nous a obligés à approximer la variable par la moyenne communale, ce qui a potentiellement introduit un aléa dans l'estimation des SIR lors du lissage.

Dans les résultats des modèles de lissage, l'association entre tabagisme et cancer de la vessie a été retrouvée chez les hommes et chez les femmes, à l'échelle de la commune comme à l'échelle de l'Iris, bien que significative uniquement chez les hommes à l'échelle communale. L'association négative apparente (statistiquement significative) entre les diagnostics de cancers du rein chez la femme et le tabagisme est un résultat contre-intuitif puisque le tabagisme est également, mais dans une moindre mesure, un facteur de risque des cancers du rein (67). Ce résultat appelle à rester critique sur l'estimation du tabagisme par le *proxy* utilisé. En effet, la mortalité par cancer du poumon n'est pas le reflet exact de la prévalence du tabagisme dans la population. Si le tabac est en effet un facteur de risque important pour ce cancer, avec une fraction attribuable de décès estimée à 19 % chez les hommes et 8 % chez les femmes en région PACA en 2015 (83), d'autres facteurs de risque sont également capturés dans ce *proxy*. Ainsi, les fractions attribuables aux expositions professionnelles dans la survenue de nouveaux cas de cancers du poumon étaient estimées à 22 % chez les hommes et 3 % chez les femmes en France en 2015 (84). Une autre exposition bien documentée du cancer du poumon est la pollution atmosphérique. Kulhánová *et al* ont estimé ainsi qu'environ 1 500 cas de cancer du poumon étaient attribuables à l'exposition aux PM<sub>2,5</sub> en France, représentant environ 4 % des nouveaux cas de cancer du poumon. Le nombre de décès par cancer du poumon relatif à l'exposition aux particules fines était estimé à 2 500 morts en 2015 soit 7,4 % des décès par ce cancer (85). Dans cette étude, la cartographie des concentrations moyennes annuelles en PM<sub>2,5</sub> en métropole a mis en évidence des concentrations particulièrement élevées sur le département des Bouches-du-Rhône.

Ce *proxy* présente également potentiellement un biais de « survie sélective ». Il est en effet possible d'imaginer que les personnes décédées d'un cancer du poumon sont aussi celles qui étaient le plus à risque de développer un cancer de vessie. Enfin, ce *proxy* intègre également une dimension d'inégalités sociales, le risque de décès par cancer du poumon est en effet lié aux conditions socio-économiques des individus. Cette dimension pourrait également expliquer l'absence de lien mis en évidence dans notre étude entre l'indice de défavorisation sociale Fdep et le cancer de vessie du fait d'une association plus forte avec le *proxy* du tabagisme.

Il serait donc intéressant d'effectuer des analyses de sensibilité sur l'influence du tabagisme dans la répartition spatiale des cancers étudiés en utilisant d'autres *proxys*, à défaut de disposer dans un avenir proche, de données sur la proportion de fumeurs en population générale. En l'absence de ces données, la mise à disposition des volumes de vente de cigarettes par commune permettrait déjà de mieux appréhender la prévalence du tabagisme. La discussion autour des limites du *proxy* du tabagisme utilisé dans notre étude va dans le sens d'un plaidoyer pour améliorer l'estimation du tabagisme en France (14) en réalisant des mesures directes avec une précision à l'échelle communale afin de pouvoir intégrer ce facteur dans des études écologiques.

### 5.5.2 Lissage spatial des risques

Pour les modèles spatiaux, le choix d'une unité spatiale approprié est important. L'unité spatiale doit ainsi être suffisamment grande pour fournir des indicateurs de santé stables et assez petite pour que les caractéristiques socio-économiques ou d'exposition soient homogènes. Le découpage administratif peut ne pas être toujours pertinent d'un point de vue épidémiologique et ne pas toujours refléter la répartition réelle des pathologies étudiées mais en pratique, l'échelle est imposée par la disponibilité des données (23). La taille et la forme des unités spatiales utilisées peuvent en effet faire varier les résultats (86,87).

Dans cette étude, la localisation des cas de cancer à l'échelle de l'Iris et la mise à disposition de la structure de la population, de la densité et de l'indice de défaveur sociale (Fdep) à cette échelle ont permis de réaliser les analyses à deux échelles géographiques différentes, la

commune et l'Iris, ce qui a permis de comparer les résultats obtenus. Néanmoins, l'indisponibilité de l'indice d'accès aux soins retenu (APL) et du *proxy* du tabagisme (nombre de décès par cancer du poumon) à l'échelle des Iris constitue une première limite a priori dans la qualité des résultats obtenus à cette échelle, en moyennant l'effet de ces covariables sur l'ensemble des Iris de la commune.

L'échelle communale est sujette à un biais écologique évident de par l'hétérogénéité potentiellement importante des contextes (densité de population, accès aux infrastructures...) au sein d'une même commune. L'échelle de l'Iris permet de diminuer ce biais en utilisant des unités plus homogènes par la définition même des Iris et par la taille réduite du territoire couvert. Le revers principal est la diminution du nombre de cas par Iris. De plus, les maladies étudiées étant relativement rares, le nombre d'Iris ne présentant aucun cas est élevé, hormis pour le cancer de la vessie chez l'homme, créant ainsi une très forte hétérogénéité locale des valeurs de SIR. Cette hétérogénéité diminue de fait artificiellement la continuité territoriale des risques, avec pour résultat de faire varier considérablement les résultats des modèles de lissage (88). Le seuil arbitraire de 80 % d'Iris présentant au moins un cas pour considérer les analyses à l'échelle des Iris pourrait être réévalué, mais il possède l'avantage de constituer un critère commun a priori.

La définition de la zone d'étude, correspondant au découpage administratif du département, limite les résultats au territoire considéré. Dans le cas présent, il existe une surincidence de cancers de la vessie dans la partie sud-est des Bouches-du-Rhône, frontalière avec le Var. Il est possible que ce cluster s'étende également aux communes voisines du département limitrophe. Dans la composante spatiale du modèle de lissage, le voisinage des unités spatiales utilisé pour estimer leur risque ne considère aucune unité en dehors de la zone d'étude. Ces effets de bords impactent potentiellement l'estimation de la valeur de SIR, et impactent la variance puisqu'elle est inversement proportionnelle au nombre de voisins (89). Dans cette étude nous n'avons pas mis en place de mesures correctives potentielles telles que l'adaptation du modèle pour les unités frontalières, ou l'extension de la zone d'étude à une ceinture de communes « tampons » (par absence de données) (89). Cette limite est décrite dans la plupart des études publiées qui utilisent des données issues de dispositif territorial ou administratif.

Le choix de la matrice de voisinage des modèles autorégressifs a également une influence importante sur les résultats (90). Dans l'étude de 2019, le choix s'était porté sur une matrice de contiguïté de type Queen (frontière commune même ponctuelle) d'après la littérature (91). Afin d'assurer la comparabilité des résultats, nous avons conservé ce choix.

La variabilité des estimations de taux dans les unités spatiales dépend de la taille de la population (ou du nombre de cas attendus proportionnel à cette population). Les estimations des SIR des unités les moins peuplées ont donc plus de chances de présenter des valeurs extrêmes, ce qui peut résulter en une mauvaise estimation de la répartition réelle du risque sur le territoire (92). Le but principal des méthodes de lissage est de diminuer ces différences de précision des estimations d'incidence, en partageant l'information apportée par les différentes unités spatiales (23,92). Les méthodes de lissage spatial prennent également en compte une autocorrélation spatiale, qui part du principe que les unités spatiales voisines partagent un profil de risque similaire suivant la première loi de la géographie de Tobler stipulant que « ... deux objets proches ont plus de chances [d'interagir] que deux objets éloignés. » (93). Les méthodes de lissage non spatial, ne prenant pas en compte la structure spatiale des données, présentent de moins bonnes performances et ont tendance à sur-lisser les particularités locales (92,94). Dans cette étude, nous avons conservé le modèle de lissage BYM qui reste un des modèles de régression spatiale bayésien le plus couramment utilisé (95). En tant que modèle de régression, il offre en plus du lissage, la possibilité d'étudier l'effet de covariables sur les SIR (96).

Néanmoins, comme confirmé en 2020 par Duncan et Mengersen, le modèle BYM a une tendance au surlissage, phénomène souligné précédemment par plusieurs auteurs (95). Les cancers, considérés étant des maladies rares, il pourrait être intéressant de se pencher sur d'autres méthodes de lissage pour détecter des surincidences locales potentiellement gommées par le modèle BYM.

D'autres méthodes de lissage spatial existent permettant de modéliser une autocorrélation spatiale et d'intégrer l'effet de facteurs de risque connus (23,96,97), (98-100). Aamodt *et al* ont comparé trois méthodes par simulations (SaTScan, modèle BYM et modèle GAM – *Generalized additive model*) et rapportaient une bonne capacité globale du BYM à rendre compte de clusters avec un RR supérieur à 1,5. Cependant, il notait que le GAM était à la fois plus sensible et moins spécifique, et avait, comme SaTScan, une meilleure capacité à détecter des clusters de forme irrégulière (98). En outre, cette modélisation a recours à une définition du voisinage simple, identique à celle utilisée dans le modèle BYM. L'utilisation de ces modèles dans l'étude des cancers reste rare et une analyse comparative sera menée sur les données de l'Observatoire. Le recours à une inférence bayésienne basée sur les INLA (*Integrated nested Laplace approximations*) pourrait également être testé. Cette méthode développée par Rue *et al* en 2009 (101,102) est utilisable dans une très grande variété de modèles existant incluant le modèle BYM (103) et les modèles GAM spatiaux (100).

D'autres limites soulevées lors de la précédente étude persistent, comme le fait que les cartographies de risque ne prennent pas en compte le temps de latence des maladies étudiées et sont réalisées à partir de la localisation des cas au moment du diagnostic, sans tenir compte des zones de résidence antérieure. De même, du fait du caractère multifactoriel des cancers, les facteurs de risques individuels, fortement corrélés à ces maladies, influencent sans doute davantage les résultats que les cofacteurs agrégés pris en compte, qui expliquent sans doute une part plus faible des cancers. Ces limites sont cependant inhérentes à ce type d'études sur des données agrégées.

### 5.5.3 Mesure de l'agrégation et détection de clusters

Les méthodes globales de clustering permettent de donner une information sur l'hétérogénéité et la tendance globale à l'agrégation sur le territoire. Néanmoins, la significativité des tests n'implique pas nécessairement l'existence d'un cluster et inversement (40). Parmi les nombreux indices et tests associés existants, nous avons choisi en 2019 d'utiliser la statistique de Potthoff-Whittinghill pour évaluer la surdispersion (hétérogénéité), ainsi que l'indice de Moran et la statistique de Tango pour évaluer la tendance à l'agrégation. Cette liste a été complétée par l'EBI qui est une version « améliorée » de l'indice de Moran dont le principal défaut est de ne pas prendre en compte les populations des unités spatiales comme dénominateurs ainsi que la forte variabilité associée aux zones de faible population. La statistique de Tango et l'EBI présentent de bonnes performances en comparaison d'autres indices (35,40,104). L'utilisation conjointe de plusieurs indices permet de mieux apprécier l'existence effective d'une hétérogénéité et/ou d'une agrégation de cas. Néanmoins, ils ne nous renseignent pas sur la localisation et le surrisque des clusters potentiels, et doivent donc être complétés par des méthodes locales de détection.

Tout comme pour les outils de détection d'agrégation globale, il existe une grande variété d'outils de détection locale répondant à diverses problématiques. Parmi les méthodes couramment utilisées, on peut citer les indicateurs locaux d'association spatiale (LISA) (105), la famille des statistiques de scan avec une fenêtre circulaire (38), elliptique (43) ou une fenêtre flexible (106). De multiples comparaisons de méthodes ont été menées sur des données simulées ou réelles, que ce soit pour réaliser une revue des méthodes ou l'introduction de nouveaux outils (40,41,44,96). Certaines méthodes répondent à des besoins de détection particuliers comme le scan à fenêtre flexible qui offre une meilleure capacité à détecter des

clusters en U (41), mais nous avons conservé la méthode sélectionnée précédemment, à savoir la statistique de scan à fenêtre elliptique, car elle présente les avantages d'avoir de bonnes performances globales, grâce à une bonne adaptabilité, à une facilité de mise en œuvre (grâce au logiciel SaTScan (48)).

L'utilisation d'une fenêtre elliptique implique de se poser la question de vouloir ou non favoriser les clusters les plus compacts par rapport aux clusters linéaires, dont la continuité territoriale n'est pas assurée (43). Nous avons fait le choix *a priori* de conserver une flexibilité de forme tout en favorisant les clusters compacts, car la répartition des données brutes et les résultats de lissage n'étaient pas en faveur de l'existence de clusters linéaires.

La principale difficulté de l'utilisation des statistiques de scan est de fixer la taille maximale de la fenêtre de détection, définie en termes de pourcentage de population du territoire incluse dans la fenêtre. Elle est usuellement fixée au maximum de 50 % lorsqu'aucune hypothèse n'est formulée sur la répartition des risques ou les objectifs de détection (42). La taille de la fenêtre maximale de détection affecte pourtant la qualité des résultats : une fenêtre trop grande détectera plus facilement un grand cluster de risque faible, alors qu'une fenêtre trop petite pourra produire plusieurs minuscules clusters n'apportant pas beaucoup plus d'information que la visualisation des résultats bruts (42,45). La solution serait d'effectuer l'analyse avec différentes tailles mais cela implique i) de répéter l'analyse en rallongeant le temps de calcul, ii) de prendre en compte la multiplicité des tests et iii) de réaliser une sélection *a posteriori* selon les résultats obtenus (44). L'optimisation de la taille maximale de clusters rapportés par l'indice de Gini permet de répondre efficacement à cette problématique en réalisant l'ensemble des analyses en parallèle et d'effectuer le choix sur un critère défini *a priori* (44). La comparaison de cette méthode avec les autres méthodes de scan classiques montre de bonnes performances de la version utilisant une fenêtre elliptique, offrant notamment une meilleure capacité de détection des clusters de forme irrégulière, en obtenant des résultats comparables à une fenêtre flexible (45).

La définition d'une limite haute des tailles maximales de fenêtre à parcourir reste néanmoins nécessaire, mais peut être effectuée uniquement sur des critères de pertinence épidémiologique des résultats souhaités, selon l'échelle notamment. Le choix de 30 % de la population au maximum à l'échelle communale et de 10 % pour l'échelle des Iris semble être un bon compromis, notamment pour pouvoir discriminer des clusters dans la ville de Marseille (43 % de la population) et dans son environnement proche.

La recherche de clusters aux deux échelles spatiales disponibles a permis de croiser l'information obtenue et de renforcer les résultats. L'échelle des Iris offre une meilleure résolution des clusters détectés mais la distribution des cas présente une plus grande hétérogénéité due aux faibles populations de certains Iris, assimilable à du bruit (107). À l'opposé, l'échelle des communes présente moins de valeurs extrêmes, mais un niveau de détail moins grand. De plus, la forme et la taille irrégulière des communes comme des Iris a un effet sur les distances entre les centroïdes des unités spatiales, et donc les unités les plus susceptibles d'être agrégées dans un cluster. Le croisement des résultats de deux différents découpages permet donc de renforcer la qualité de l'information obtenue.

Une autre limite inhérente à toute analyse spatiale reste les effets de bord qui résultent principalement de la censure spatiale (i.e. l'absence de prise en compte des cas, populations et cofacteurs au-delà du territoire de l'étude) (89). Ces effets diminuent la probabilité de détecter des clusters aux bords, ce qui implique par ailleurs que le risque de faux positifs aux frontières du territoire est faible (56,108). L'utilisation d'un découpage plus cohérent et mieux réparti comme les Iris par rapport aux communes, semble réduire les effets de bords (56,109). Plusieurs méthodes ont été proposées pour diminuer les effets de bords : i) modifier les outils de façon à prendre en compte la proximité des unités spatiales aux bords, ii) étendre la zone d'analyse au-delà de la zone d'intérêt, ou iii) simuler des estimations des territoires frontaliers

ou les considérer comme des valeurs manquantes (89). Dans cette étude, l'importante zone côtière (maritime et étang de Berre), impliquant une absence réelle de voisinage, limite la pertinence des méthodes impliquant une simulation de voisinage mais l'existence d'un cluster de cancers de la vessie chez les hommes à la frontière avec le Var pourrait être une raison de les explorer.

## 6 CONCLUSIONS

Les nouvelles analyses présentées dans ce rapport viennent compléter les résultats du rapport précédent publié en 2019 (6), en ajoutant deux années d'incidence supplémentaires (2017 et 2018). La réalisation des analyses spatiales à une échelle géographique plus fine (Iris) était recommandée dans le précédent rapport et a pu être réalisée pour les cancers de vessie chez les hommes, mais pas chez les femmes ni pour les deux autres cancers, du fait d'un nombre de cas limité à cette échelle. Les résultats confirment l'existence d'une sur-incidence de cancers de vessie à l'échelle du département, plus marquée sur les communes densément peuplées et où un tabagisme important est estimé, suggérant l'existence d'expositions environnementales et socio-économiques communes à ces zones.

La prise en compte du facteur de risque principal du cancer de vessie dans ces nouvelles analyses a montré que le proxy du tabagisme influence effectivement la répartition spatiale des cancers de vessie mais qu'il n'explique pas, à lui seul, les différences d'incidence observées. La densité de population, également identifiée dans la littérature comme facteur associé à l'incidence des cancers de vessie, est liée positivement à la répartition spatiale de ce cancer dans notre étude.

L'exposition à la pollution atmosphérique pourrait également être considéré comme un facteur de risque du cancer de la vessie. La réalisation d'une étude écologique géographique permettrait de vérifier cette hypothèse. Cette approche pourrait être développée soit à partir d'indicateurs traceurs de la pollution atmosphérique urbaine, comme le dioxyde d'azote et les particules fines, mais aussi de polluants traceurs de la pollution industrielle, soit à partir de la construction d'un indicateur de multiexpositions. Ce type d'indicateur a récemment été utilisé par l'équipe du centre Léon Bérard à Lyon dans l'étude Xénair visant à mettre en évidence un lien entre pollution atmosphérique et risque de cancer du sein (110).

Santé publique France envisage cette étude écologique pour mettre en lien les données recueillies dans REVELA13 avec des données de pollution atmosphérique, en s'appuyant sur une équipe de recherche locale.

## 7 Références bibliographiques

1. Inserm(dir). Expertise Inserm : Cancer et environnement. Collect Expert Collect Inserm [Internet]. 2008 [cité 28 juin 2022]; Disponible sur: <https://www.ipubli.inserm.fr/handle/10608/102>
2. Le Moal J, Eilstein D, Straif K, Ledrans M. Cancers prioritaires à surveiller et étudier en lien avec l'environnement. [Internet]. 2006 juill p. 12. (Santé publique France). Disponible sur: <https://www.santepubliquefrance.fr/docs/cancers-prioritaires-a-surveiller-et-etudier-en-lien-avec-l-environnement>
3. Pascal L, Iwatsubo Y, Lasalle JL, Charlet F. Investigation d'une suspicion d'agrégats de cancers du rein au niveau de la plateforme pétrochimique de Lavéra, Bouches-du-Rhône [Internet]. 2010 déc p. 23. Disponible sur: <https://www.santepubliquefrance.fr/regions/provence-alpes-cote-d-azur-et-corse/documents/rapport-synthese/2010/investigation-d-une-suspicion-d-agregat-de-cancers-du-rein-au-niveau-de-la-plateforme-petrochimique-de-lavera-bouches-du-rhone-1994-2007>
4. Pascal L, Stempfelet M, Gorla S, Lasalle JL, Pascal M, Declercq C, *et al.* Pollution atmosphérique et hospitalisations pour pathologies cardio-vasculaires et respiratoires, et pour cancers dans le secteur de l'étang de Berre, 2004-2007 [Internet]. 2011 juill p. 68. (Santé publique France). Disponible sur: <https://www.santepubliquefrance.fr/regions/provence-alpes-cote-d-azur-et-corse/documents/rapport-synthese/2011/pollution-atmospherique-et-hospitalisations-pour-pathologies-cardio-vasculaires-et-respiratoires-et-pour-cancers-dans-le-secteur-de-l-etang-de-berre>
5. Pascal L, Mantey K. Observatoire des cancers du rein, de la vessie et des leucémies aiguës (Revela13) chez les adultes des Bouches-du-Rhône. Analyse de l'exhaustivité, 2013-2014. [Internet]. 2018 sept p. 72. (Santé publique France). Disponible sur: <https://www.santepubliquefrance.fr/regions/provence-alpes-cote-d-azur-et-corse/documents/rapport-synthese/2018/observatoire-des-cancers-du-rein-de-la-vessie-et-des-leucemies-aigues-revela13-chez-les-adultes-des-bouches-du-rhone-analyse-de-l-exhaustivite>
6. Mantey K, Pascal L. Observatoire des cancers du rein, de la vessie et des leucémies aiguës chez l'adulte dans le département des Bouches-du-Rhône (REVELA13) : Analyses spatiales, 2013-2016 [Internet]. 2019 juill [cité 1 févr 2022] p. 80. (Santé publique France). Disponible sur: <https://www.santepubliquefrance.fr/regions/provence-alpes-cote-d-azur-et-corse/documents/rapport-synthese/2019/observatoire-des-cancers-du-rein-de-la-vessie-et-des-leucemies-aigues-chez-l-adulte-dans-le-departement-des-bouches-du-rhone-revela13-analyses>
7. Insee. La grille communale de densité à 4 niveaux [Internet]. 2022 [cité 15 juin 2022]. Disponible sur: <https://www.insee.fr/fr/information/2114627>
8. Insee. Définition - IRIS [Internet]. 2016 [cité 15 juin 2022]. Disponible sur: <https://www.insee.fr/fr/metadonnees/definition/c1523>
9. Bryere J, Dejardin O, Launay L, Colonna M, Grosclaude P, Launoy G ; Réseau français des registres des cancers (Francim). Bull Epidemiol Hebd. 2017 [cité 1<sup>er</sup> juill 2022]. Environnement socioéconomique et incidence des cancers en France. Disponible sur: [http://beh.santepubliquefrance.fr/beh/2017/4/2017\\_4\\_1.html](http://beh.santepubliquefrance.fr/beh/2017/4/2017_4_1.html)

10. Gorza M, Eilstein D. Outils élaborés dans la cadre du programme « Inégalités sociales de santé », 2013-2015 [Internet]. 2018 p. 108. (Saint-Maurice: Santé publique France). Disponible sur:  
[https://www.santepubliquefrance.fr/content/download/225070/document\\_file/41970\\_spf0000343.pdf](https://www.santepubliquefrance.fr/content/download/225070/document_file/41970_spf0000343.pdf)
11. Rey G, Rican S, Jouglu E. Mesure des inégalités par causes de décès. Approche écologique à l'aide d'un indice de désavantage social. BEH 2011 [Internet]. 2011;(8-9):87-90. Disponible sur:  
[https://www.santepubliquefrance.fr/content/download/173137/document\\_file/32821\\_9264.pdf](https://www.santepubliquefrance.fr/content/download/173137/document_file/32821_9264.pdf)
12. Insee, CépiDc-Inserm. Indicateurs écologiques du niveau socio-économique [Internet]. [cité 15 juin 2022]. Disponible sur:  
<https://www.cepidc.inserm.fr/documentation/indicateurs-ecologiques-du-niveau-socio-economique>
13. Direction de la recherche, des études, de l'évaluation et des statistiques. L'accessibilité potentielle localisée (APL) [Internet]. [cité 2 mars 2022]. Disponible sur:  
[https://data.drees.solidarites-sante.gouv.fr/explore/dataset/530\\_l-accessibilite-potentielle-localisee-apl/information/](https://data.drees.solidarites-sante.gouv.fr/explore/dataset/530_l-accessibilite-potentielle-localisee-apl/information/)
14. Pasquereau A, Andler R, Guignard R, Richard JB, Arwidson P, Nguyen-Thanh V, *et al.* La consommation de tabac en France : premiers résultats du Baromètre santé 2017. Bull Épidémiologique Hebd [Internet]. 2018;(14-15):265-73. Disponible sur: [http://invs.santepubliquefrance.fr/beh/2018/14-15/2018\\_14-15\\_1.html](http://invs.santepubliquefrance.fr/beh/2018/14-15/2018_14-15_1.html)
15. Lopez-Abente G, Aragonés N, Ramis R, Hernandez-Barrera V, Perez-Gomez B, Escolar-Pujolar A, *et al.* Municipal distribution of bladder cancer mortality in Spain: Possible role of mining and industry. BMC Public Health [Internet]. déc 2006 [cité 18 nov 2021];6(1):17. Disponible sur:  
<https://bmcpublihealth.biomedcentral.com/articles/10.1186/1471-2458-6-17>
16. Fazzo L, Carere M, Tisano F, Bruno C, Cernigliaro A, Cicero MR, *et al.* Cancer incidence in Priolo, Sicily: a spatial approach for estimation of industrial air pollution impact. Geospatial Health. 18 avr 2016;11(1):320.
17. Verhoeven RHA, Louwman MWJ, Buntinx F, Botterweck AM, Lousbergh D, Faes C, *et al.* Variation in cancer incidence in northeastern Belgium and southeastern Netherlands seems unrelated to cadmium emission of zinc smelters. Eur J Cancer Prev Off J Eur Cancer Prev Organ ECP. nov 2011;20(6):549-55.
18. Rican S. La cartographie des données épidémiologiques. Les principales méthodes de discrétisation et leur importance dans la représentation cartographique. Cah d'études Rech Francoph Santé. 1999;8(6):461-70.
19. Ancelle T. Statistique épidémiologie. Paris Éditions Maloine. 2002.
20. Defossez G, Le Guyader-Peyrou S, Uhry Z, Grosclaude P, Colonna M, *et al.* Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. Volume 1 – Tumeurs solides [Internet]. Saint-Maurice (Fra): Santé publique France; 2019 p. 372. Disponible sur:  
<https://www.santepubliquefrance.fr/maladies-et-traumatismes/cancers/cancer-du-sein/documents/rapport-synthese/estimations-nationales-de-l-incidence-et-de-la-mortalite-par-cancer-en-france-metropolitaine-entre-1990-et-2018-volume-1-tumeurs-solides-etud>

21. Le Guyader-Peyrou S, Defossez G, Uhry Z, Mounier M, Cornet E, Dantony E *et al.* Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. Volume 2 – Hémopathies malignes. Étude à partir des registres des cancers du réseau Francim. [Internet]. Saint-Maurice (Fra): Santé publique France; 2019 p. 169. Disponible sur: <https://www.santepubliquefrance.fr/docs/estimations-nationales-de-l-incidence-et-de-la-mortalite-par-cancer-en-france-metropolitaine-entre-1990-et-2018-volume-2-hemopathies-malignes>
22. International Agency for Research on Cancer. A New Measure of Age Standardized Incidence, the Cumulative Rate. In: Cancer incidence in five continents 3. Lyon: Waterhouse JAH, Muir CS, Correa P, Powell J; 1976. p. 443-52. (IARC scientific publications).
23. Gorla S, Stempfelet M, de Crouy P. Introduction aux statistiques spatiales et aux systèmes d'information géographique en santé environnement. [Internet]. Saint-Maurice (Fra): Institut de veille sanitaire; 2011 [cité 17 juin 2022] p. 65. Disponible sur: <https://www.santepubliquefrance.fr/notices/introduction-aux-statistiques-spatiales-et-aux-systemes-d-information-geographique-en-sante-environnement.-application-aux-etudes-ecologiques.-resu>
24. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math* [Internet]. 1<sup>er</sup> mars 1991 [cité 22 févr 2022];43(1):1-20. Disponible sur: <https://doi.org/10.1007/BF00116466>
25. Besag J. Spatial Interaction and the Statistical Analysis of Lattice Systems. *J R Stat Soc Ser B Methodol* [Internet]. janv 1974 [cité 19 juin 2022];36(2):192-225. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1111/j.2517-6161.1974.tb00999.x>
26. Schmidt AM, Nobre WS. Conditional Autoregressive (CAR) Model. In: Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels JL, éditeurs. *Wiley StatsRef: Statistics Reference Online* [Internet]. 1<sup>re</sup> éd. Wiley; 2018 [cité 17 juin 2022]. p. 1-11. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/9781118445112.stat08048>
27. Colonna M. Influence des paramètres a priori dans l'estimation bayésienne de risques relatifs. *Analyse spatiale du cancer de la vessie dans l'agglomération grenobloise. Rev d'Épidémiologie Santé Publique* [Internet]. déc 2006 [cité 9 nov 2022];54(6):529-42. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0398762006767529>
28. Colonna M, Sauleau EA. How to interpret and choose a Bayesian spatial model and a Poisson regression model in the context of describing small area cancer risks variations. *Rev d'Épidémiologie Santé Publique* [Internet]. 1 déc 2013 [cité 22 nov 2021];61(6):559-67. Disponible sur: <https://www.sciencedirect.com/science/article/pii/S0398762013010493>
29. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol* [Internet]. oct 2002 [cité 9 nov 2022];64(4):583-639. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1111/1467-9868.00353>
30. Moran PAP. Notes on continuous stochastic phenomena. *Biometrika* [Internet]. 1950 [cité 19 juin 2022];37(1-2):17-23. Disponible sur: <https://academic.oup.com/biomet/article-lookup/doi/10.1093/biomet/37.1-2.17>
31. Salel, C, Catelinois, O, Cariou, M, Billot-Grasset A, Chatignoux, E. Estimations régionales et départementales d'incidence et de mortalité par cancers en France, 2007-2016. PACA [Internet]. 2019 janv p. 167. (Santé publique France). Disponible sur: <https://www.santepubliquefrance.fr/>

32. R Development Core Team. R : a language and environment for statistical computing: reference index [Internet]. Vienna: R Foundation for Statistical Computing; 2010 [cité 19 juin 2022]. Disponible sur: <http://www.polsci.wvu.edu/duval/PS603/Notes/R/fullrefman.pdf>
33. Lee D. CARBayes version 4.7: An R Package for Spatial Areal Unit Modelling with Conditional Autoregressive Priors.29.
34. Potthoff RF, Whittinghill M. Testing for homogeneity. II. The Poisson distribution. *Biometrika*. juin 1966;53(1):183-90.
35. Assunção RM, Reis EA. A new proposal to adjust Moran's I for population density. *Stat Med* [Internet]. 30 août 1999 [cité 19 juin 2022];18(16):2147-62. Disponible sur: [https://doi.org/10.1002/\(sici\)1097-0258\(19990830\)18:16%3C2147::aid-sim179%3E3.0.co;2-j](https://doi.org/10.1002/(sici)1097-0258(19990830)18:16%3C2147::aid-sim179%3E3.0.co;2-j)
36. Tango T. A class of tests for detecting 'general' and 'focused' clustering of rare diseases. *Stat Med* [Internet]. 15 nov 1995 [cité 19 juin 2022];14(21-22):2323-34. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/sim.4780142105>
37. Kulldorff M, Nagarwalla N. Spatial disease clusters: Detection and inference. *Stat Med* [Internet]. 30 avr 1995 [cité 19 juin 2022];14(8):799-810. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/sim.4780140809>
38. Kulldorff M. A spatial scan statistic. *Commun Stat - Theory Methods* [Internet]. janv 1997 [cité 14 févr 2022];26(6):1481-96. Disponible sur: <http://www.tandfonline.com/doi/abs/10.1080/03610929708831995>
39. Kulldorff M. Spatial Scan Statistics: Models, Calculations, and Applications. In: Glaz J, Balakrishnan N, éditeurs. *Scan Statistics and Applications* [Internet]. Boston, MA: Birkhäuser Boston; 1999 [cité 14 févr 2022]. p. 303-22. Disponible sur: [http://link.springer.com/10.1007/978-1-4612-1578-3\\_14](http://link.springer.com/10.1007/978-1-4612-1578-3_14)
40. Huang L, Pickle LW, Das B. Evaluating spatial methods for investigating global clustering and cluster detection of cancer cases. *Stat Med* [Internet]. 10 nov 2008 [cité 22 oct 2021];27(25):5111-42. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/sim.3342>
41. 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* [Internet]. 2011 [cité 19 juin 2022];10(1):53. Disponible sur: <http://ij-healthgeographics.biomedcentral.com/articles/10.1186/1476-072X-10-53>
42. Ribeiro SHR, Costa MA. Optimal selection of the spatial scan parameters for cluster detection: A simulation study. *Spat Spatio-Temporal Epidemiol* [Internet]. juin 2012 [cité 19 juin 2022];3(2):107-20. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S1877584512000330>
43. Kulldorff M, Huang L, Pickle L, Duczmal L. An elliptic spatial scan statistic: ELLIPTIC SPATIAL SCAN STATISTIC. *Stat Med* [Internet]. 30 nov 2006 [cité 14 févr 2022];25(22):3929-43. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/sim.2490>
44. Han J, Zhu L, Kulldorff M, Hostovich S, Stinchcomb DG, Tatalovich Z, *et al*. Using Gini coefficient to determining optimal cluster reporting sizes for spatial scan statistics. *Int J Health Geogr* [Internet]. déc 2016 [cité 30 mai 2022];15(1):27. Disponible sur: <http://ij-healthgeographics.biomedcentral.com/articles/10.1186/s12942-016-0056-6>
45. Kim J, Jung I. Evaluation of the Gini Coefficient in Spatial Scan Statistics for Detecting Irregularly Shaped Clusters. Meliker J, éditeur. *PLOS ONE* [Internet]. 27 janv 2017 [cité 19 juin 2022];12(1):e0170736. Disponible sur: <https://dx.plos.org/10.1371/journal.pone.0170736>

46. Gini C. Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche. [Fasc. I.] [Internet]. Tipogr. di P. Cuppini; 1912. (Studi economico-giuridici pubblicati per cura della facoltà di Giurisprudenza della R. Università di Cagliari). Disponible sur: <https://books.google.fr/books?id=fqjaBPMxB9kC>
47. Lorenz MO. Methods of Measuring the Concentration of Wealth. *Publ Am Stat Assoc* [Internet]. juin 1905 [cité 18 nov 2022];9(70):209. Disponible sur: <https://www.jstor.org/stable/2276207?origin=crossref>
48. Kulldorff M. and Information Management Services. Software for the spatial and spacetime scan statistics - Inc. SaTScan™ [Internet]. 2021. Disponible sur: <https://www.satscan.org/>
49. Dobruch J, Daneshmand S, Fisch M, Lotan Y, Noon AP, Resnick MJ, *et al.* Gender and Bladder Cancer: A Collaborative Review of Etiology, Biology, and Outcomes. *Eur Urol* [Internet]. févr 2016 [cité 21 juill 2022];69(2):300-10. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0302283815007824>
50. Hyldgaard JM, Jensen JB. The Inequality of Females in Bladder Cancer. *APMIS* [Internet]. déc 2021 [cité 21 juill 2022];129(12):694-9. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1111/apm.13183>
51. Cohn JA, Vekhter B, Lyttle C, Steinberg GD, Large MC. Sex disparities in diagnosis of bladder cancer after initial presentation with hematuria: A nationwide claims-based investigation: Delay in Bladder Cancer Diagnosis. *Cancer* [Internet]. 15 févr 2014 [cité 21 juill 2022];120(4):555-61. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/cncr.28416>
52. Mancini M, Righetto M, Baggio G. Spotlight on gender-specific disparities in bladder cancer. *Urol J* [Internet]. août 2020 [cité 21 juill 2022];87(3):103-14. Disponible sur: <http://journals.sagepub.com/doi/10.1177/0391560319887327>
53. Han DS, Zhou W, Seigne JD, Lynch KE, Schroeck FR. Geographic Variation in Cystoscopy Rates for Suspected Bladder Cancer between Female and Male Medicare Beneficiaries. *Urology* [Internet]. déc 2018 [cité 8 sept 2022];122:83-8. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0090429518308355>
54. Saint-Jacques N, Lee JSW, Brown P, Stafford J, Parker L, Dummer TJB. Small-area spatio-temporal analyses of bladder and kidney cancer risk in Nova Scotia, Canada. *BMC Public Health* [Internet]. 19 févr 2016 [cité 8 sept 2022];16(1):175. Disponible sur: <https://doi.org/10.1186/s12889-016-2767-9>
55. Cao B, Hill C, Bonaldi C, León ME, Menvielle G, Arwidson P, *et al.* Cancers attributable to tobacco smoking in France in 2015. *Eur J Public Health* [Internet]. 1 août 2018 [cité 22 juill 2022];28(4):707-12. Disponible sur: <https://academic.oup.com/eurpub/article/28/4/707/4993388>
56. Goungounga JA, Gaudart J, Colonna M, Giorgi R. Impact of socioeconomic inequalities on geographic disparities in cancer incidence: comparison of methods for spatial disease mapping. *BMC Med Res Methodol* [Internet]. 2016 [cité 11 févr 2022];16(1):136. Disponible sur: <https://hal.archives-ouvertes.fr/hal-01873137>
57. Bako G, Dewar R, Hanson J, Hill G. Population density as an indicator of urban-rural differences in cancer incidence, Alberta, Canada, 1969-73. *Can J Public Health Rev Can Sante Publique*. avr 1984;75(2):152-6.
58. Nasca PC, Burnett WS, Greenwald P, Brennan K, Wolfgang P, Carlton K. Population density as an indicator of urban-rural differences in cancer incidence, Upstate New-York, 1968–19721. *Am J Epidemiol* [Internet]. sept 1980 [cité 22 juill 2022];112(3):362-75. Disponible sur: <https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a113002>

59. Radespiel-Tröger M, Geiss K, Twardella D, Maier W, Meyer M. Cancer incidence in urban, rural, and densely populated districts close to core cities in Bavaria, Germany. *Int Arch Occup Environ Health* [Internet]. févr 2018 [cité 22 juill 2022];91(2):155-74. Disponible sur: <http://link.springer.com/10.1007/s00420-017-1266-3>
60. Castano-Vinyals G, Cantor KP, Malats N, Tardon A, Garcia-Closas R, Serra C, *et al.* Air pollution and risk of urinary bladder cancer in a case-control study in Spain. *Occup Environ Med* [Internet]. 1 janv 2008 [cité 3 févr 2022];65(1):56-60. Disponible sur: <https://oem.bmj.com/lookup/doi/10.1136/oem.2007.034348>
61. Zare Sakhvidi MJ, Lequy E, Goldberg M, Jacquemin B. Air pollution exposure and bladder, kidney and urinary tract cancer risk: A systematic review. *Environ Pollut* [Internet]. déc 2020 [cité 3 févr 2022];267:115328. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0269749120360164>
62. Cumberbatch MGK, Jubber I, Black PC, Esperto F, Figueroa JD, Kamat AM, *et al.* Epidemiology of Bladder Cancer: A Systematic Review and Contemporary Update of Risk Factors in 2018. *Eur Urol* [Internet]. déc 2018 [cité 4 févr 2022];74(6):784-95. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0302283818306511>
63. Vieira V, Webster T, Weinberg J, Aschengrau A. Spatial analysis of bladder, kidney, and pancreatic cancer on upper Cape Cod: an application of generalized additive models to case-control data. *Environ Health* [Internet]. 10 févr 2009 [cité 8 sept 2022];8(1):3. Disponible sur: <https://doi.org/10.1186/1476-069X-8-3>
64. Fazzo L, De Santis M, Mitis F, Benedetti M, Martuzzi M, Comba P, *et al.* Ecological studies of cancer incidence in an area interested by dumping waste sites in Campania (Italy). *Ann Ist Super Sanita.* 2011;47(2):181-91.
65. Leiser CL, Taddie M, Hemmert R, Richards Steed R, VanDerslice JA, Henry K, *et al.* Spatial clusters of cancer incidence: analyzing 1940 census data linked to 1966–2017 cancer records. *Cancer Causes Control* [Internet]. 1 juill 2020 [cité 8 sept 2022];31(7):609-15. Disponible sur: <https://doi.org/10.1007/s10552-020-01302-3>
66. Fazzo L, Belli S, Minichilli F, Mitis F, Santoro M, Martina L, *et al.* Cluster analysis of mortality and malformations in the Provinces of Naples and Caserta (Campania Region). *Ann Ist Super Sanita.* 2008;44(1):99-111.
67. Scelo G, Larose TL. Epidemiology and Risk Factors for Kidney Cancer. *J Clin Oncol* [Internet]. 20 déc 2018 [cité 6 mai 2022];36(36):3574-81. Disponible sur: <https://ascopubs.org/doi/10.1200/JCO.2018.79.1905>
68. Capitanio U, Bensalah K, Bex A, Boorjian SA, Bray F, Coleman J, *et al.* Epidemiology of Renal Cell Carcinoma. *Eur Urol* [Internet]. janv 2019 [cité 29 juill 2022];75(1):74-84. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S030228381830633X>
69. Cumberbatch MG, Rota M, Catto JWF, La Vecchia C. The Role of Tobacco Smoke in Bladder and Kidney Carcinogenesis: A Comparison of Exposures and Meta-analysis of Incidence and Mortality Risks. *Eur Urol* [Internet]. sept 2016 [cité 29 juill 2022];70(3):458-66. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0302283815005485>
70. Eriksen KT, Petersen A, Poulsen AH, Deltour I, Raaschou-Nielsen O. Social inequality and incidence of and survival from cancers of the kidney and urinary bladder in a population-based study in Denmark, 1994–2003. *Eur J Cancer* [Internet]. sept 2008 [cité 9 sept 2022];44(14):2030-42. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0959804908004668>
71. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. Longo DL, éditeur. *N Engl J Med* [Internet]. 17 sept 2015 [cité 8 sept 2022];373(12):1136-52. Disponible sur: <http://www.nejm.org/doi/10.1056/NEJMra1406184>

72. Kulsoom B, Shamsi TS, Ahmed N, Hasnain SN. Clinical presentation of acute myeloid leukaemia - A decade-long institutional follow-up. *JPMA J Pak Med Assoc.* déc 2017;67(12):1837-42.
73. Roman E, Smith A, Appleton S, Crouch S, Kelly R, Kinsey S, *et al.* Myeloid malignancies in the real-world: Occurrence, progression and survival in the UK's population-based Haematological Malignancy Research Network 2004–15. *Cancer Epidemiol* [Internet]. juin 2016 [cité 8 sept 2022];42:186-98. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S1877782116300364>
74. Sasaki K, Ravandi F, Kadia TM, DiNardo CD, Short NJ, Borthakur G, *et al.* De novo acute myeloid leukemia: A population-based study of outcome in the United States based on the Surveillance, Epidemiology, and End Results (SEER) database, 1980 to 2017. *Cancer* [Internet]. 15 juin 2021 [cité 8 sept 2022];127(12):2049-61. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1002/cncr.33458>
75. Maynadié M, Manivet I, Mounier M, Mugneret F, Bailly F, Caillot D, *et al.* Trends in Incidence and Survival of Myeloid Malignancies Since 1980, in the Côte D'or Department, Burgundy, France. *Blood* [Internet]. 20 nov 2009 [cité 16 sept 2022];114(22):3122-3122. Disponible sur: <https://ashpublications.org/blood/article/114/22/3122/132742/Trends-in-Incidence-and-Survival-of-Myeloid>
76. Contreras ZA, Hansen J, Ritz B, Olsen J, Yu F, Heck JE. Parental age and childhood cancer risk: A Danish population-based registry study. *Cancer Epidemiol* [Internet]. août 2017 [cité 13 sept 2022];49:202-15. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S1877782117300991>
77. Gaubert G. Analyse géo-épidémiologique des données d'hospitalisations pour Covid-19 en PACA [Internet]. 2021 [cité 8 août 2022]. Disponible sur: <https://dumas.ccsd.cnrs.fr/dumas-03432940>
78. Meijer M, Bloomfield K, Engholm G. Neighbourhoods matter too: the association between neighbourhood socioeconomic position, population density and breast, prostate and lung cancer incidence in Denmark between 2004 and 2008. *J Epidemiol Community Health* [Internet]. 1 janv 2013 [cité 29 juill 2022];67(1):6-13. Disponible sur: <https://jech.bmj.com/content/67/1/6>
79. Hagedoorn P, Vandenheede H, Vanthomme K, Gadeyne S. Socioeconomic position, population density and site-specific cancer mortality: A multilevel analysis of Belgian adults, 2001–2011. *Int J Cancer* [Internet]. 2018 [cité 29 juill 2022];142(1):23-35. Disponible sur: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.31031>
80. Wan N, Zhan FB, Lu Y, Tiefenbacher JP. Access to healthcare and disparities in colorectal cancer survival in Texas. *Health Place* [Internet]. 1 mars 2012 [cité 29 juill 2022];18(2):321-9. Disponible sur: <https://www.sciencedirect.com/science/article/pii/S1353829211002048>
81. Brown S, Castelli M, Hunter DJ, Erskine J, Vedsted P, Foot C, *et al.* How might healthcare systems influence speed of cancer diagnosis: A narrative review. *Soc Sci Med* [Internet]. 1 sept 2014 [cité 29 juill 2022];116:56-63. Disponible sur: <https://www.sciencedirect.com/science/article/pii/S0277953614003980>
82. Fournel I, Bourredjem A, Sauleau EA, Cottet V, Dejardin O, Bouvier AM, *et al.* Small-area geographic and socioeconomic inequalities in colorectal tumour detection in France. *Eur J Cancer Prev.* 1 juill 2016;25(4):269-74.
83. Bulletin de santé publique. TABAC en région PACA. Santé publique France; 2021 févr.

84. Marant Micallef C, Shield KD, Vignat J, Baldi I, Charbotel B, Fervers B, *et al.* Cancers in France in 2015 attributable to occupational exposures. *Int J Hyg Environ Health* [Internet]. janv 2019 [cité 4 juill 2022];222(1):22-9. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S1438463918303262>
85. Kulhánová I, Morelli X, Le Tertre A, Loomis D, Charbotel B, Medina S, *et al.* The fraction of lung cancer incidence attributable to fine particulate air pollution in France: Impact of spatial resolution of air pollution models. *Environ Int* [Internet]. déc 2018 [cité 5 août 2022];121:1079-86. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0160412018312236>
86. Openshaw S. The modifiable areal unit problem. In: *Classification using information statistics* [Internet]. Norwich [Norfolk]: Geo Books; 1983. (Concepts and techniques in modern geography). Disponible sur: <https://www.uio.no/studier/emner/sv/iss/SGO9010/openshaw1983.pdf>
87. Fotheringham AS, Wong DWS. The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environ Plan Econ Space* [Internet]. 1 juill 1991 [cité 28 juill 2022];23(7):1025-44. Disponible sur: <https://doi.org/10.1068/a231025>
88. Boscoe FP, Ward MH, Reynolds P. Current practices in spatial analysis of cancer data: data characteristics and data sources for geographic studies of cancer. *Int J Health Geogr* [Internet]. 1 déc 2004 [cité 28 juill 2022];3(1):28. Disponible sur: <https://doi.org/10.1186/1476-072X-3-28>
89. Lawson A. *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. 2021.
90. Earnest A, Morgan G, Mengersen K, Ryan L, Summerhayes R, Beard J. Evaluating the effect of neighbourhood weight matrices on smoothing properties of Conditional Autoregressive (CAR) models. *Int J Health Geogr* [Internet]. 29 nov 2007 [cité 28 juill 2022];6:54. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2242788/>
91. Duncan EW, White NM, Mengersen K. Spatial smoothing in Bayesian models: a comparison of weights matrix specifications and their impact on inference. *Int J Health Geogr* [Internet]. 16 déc 2017 [cité 28 juill 2022];16(1):47. Disponible sur: <https://doi.org/10.1186/s12942-017-0120-x>
92. Waller LA, Gotway CA. *Applied spatial statistics for public health data*. Hoboken, N.J.: John Wiley & Sons; 2004. 494 p. (Wiley series in probability and statistics).
93. Tobler WR. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ Geogr* [Internet]. 1 juin 1970 [cité 28 juill 2022];46(sup1):234-40. Disponible sur: <https://www.tandfonline.com/doi/abs/10.2307/143141>
94. Lemke D, Mattauch V, Heidinger O, Pebesma E, Hense HW. Detecting cancer clusters in a regional population with local cluster tests and Bayesian smoothing methods: a simulation study. *Int J Health Geogr* [Internet]. 7 déc 2013 [cité 28 juill 2022];12(1):54. Disponible sur: <https://doi.org/10.1186/1476-072X-12-54>
95. Duncan EW, Mengersen KL. Comparing Bayesian spatial models: Goodness-of-smoothing criteria for assessing under- and over-smoothing. *PLOS ONE* [Internet]. 20 mai 2020 [cité 28 juill 2022];15(5):e0233019. Disponible sur: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0233019>
96. Kanankege KST, Alvarez J, Zhang L, Perez AM. An Introductory Framework for Choosing Spatiotemporal Analytical Tools in Population-Level Eco-Epidemiological Research. *Front Vet Sci* [Internet]. 2020 [cité 28 juill 2022];7. Disponible sur: <https://www.frontiersin.org/articles/10.3389/fvets.2020.00339>

97. Auchincloss AH, Gebreab SY, Mair C, Diez Roux AV. A Review of Spatial Methods in Epidemiology, 2000–2010. *Annu Rev Public Health* [Internet]. 21 avr 2012 [cité 17 juin 2022];33(1):107-22. Disponible sur: <https://www.annualreviews.org/doi/10.1146/annurev-publhealth-031811-124655>
98. Aamodt G, Samuelsen SO, Skrondal A. A simulation study of three methods for detecting disease clusters. *Int J Health Geogr*. 12 avr 2006;5:15.
99. Hawinkel S, De Meyer S, Maere S. Spatial Regression Models for Field Trials: A Comparative Study and New Ideas. *Front Plant Sci*. 2022;13:858711.
100. Guindo A, Sagara I, Ouedraogo B, Sallah K, Assadou MH, Healy S, *et al*. Spatial heterogeneity of environmental risk in randomized prevention trials: consequences and modeling. *BMC Med Res Methodol* [Internet]. 15 juill 2019 [cité 28 juill 2022];19(1):149. Disponible sur: <https://doi.org/10.1186/s12874-019-0759-z>
101. Fong Y, Rue H, Wakefield J. Bayesian inference for generalized linear mixed models. *Biostatistics* [Internet]. 1 juill 2010 [cité 28 juill 2022];11(3):397-412. Disponible sur: <https://doi.org/10.1093/biostatistics/kxp053>
102. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Ser B Stat Methodol* [Internet]. avr 2009 [cité 28 nov 2022];71(2):319-92. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9868.2008.00700.x>
103. Martino S, Riebler A. Integrated Nested Laplace Approximations (INLA) [Internet]. arXiv; 2019 [cité 28 juill 2022]. Disponible sur: <http://arxiv.org/abs/1907.01248>
104. Jackson MC, Huang L, Luo J, Hachey M, Feuer E. Comparison of tests for spatial heterogeneity on data with global clustering patterns and outliers. *Int J Health Geogr* [Internet]. 12 oct 2009 [cité 28 juill 2022];8(1):55. Disponible sur: <https://doi.org/10.1186/1476-072X-8-55>
105. Anselin L. Local Indicators of Spatial Association–ISA. *Geogr Anal*. 3 sept 2010;27:93-115.
106. Tango T, Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. *Int J Health Geogr* [Internet]. 18 mai 2005 [cité 29 juill 2022];4(1):11. Disponible sur: <https://doi.org/10.1186/1476-072X-4-11>
107. Oshan TM, Wolf LJ, Sachdeva M, Bardin S, Fotheringham AS. A scoping review on the multiplicity of scale in spatial analysis. *J Geogr Syst* [Internet]. juill 2022 [cité 28 nov 2022];24(3):293-324. Disponible sur: <https://link.springer.com/10.1007/s10109-022-00384-8>
108. Guttmann A, Li X, Gaudart J, Gérard Y, Demongeot J, Boire JY, *et al*. Spatial heterogeneity of type I error for local cluster detection tests. *Int J Health Geogr* [Internet]. 27 mai 2014 [cité 29 juill 2022];13(1):15. Disponible sur: <https://doi.org/10.1186/1476-072X-13-15>
109. Tran A, Gardon J, Weber S, Polidori L. Mapping disease incidence in suburban areas using remotely sensed data. *Am J Epidemiol*. 1 oct 2002;156(7):662-8.
110. Grassot, L, Coudon, T, Deygas, F, Dubuis, M, Amadou, A, Couvidat, F, *et al*. Vers une étude de la multiexposition : Analyse de la corrélation spatio-temporelle entre polluants de l'air [Internet]. Congrès de la société francophone de santé environnement; 2020 oct [cité 1 févr 2022]. Disponible sur: <https://www.sfse.org/congres/presentations/11>

111. Link WA, Eaton MJ. On thinning of chains in MCMC. *Methods Ecol Evol* [Internet]. févr 2012 [cité 8 nov 2022];3(1):112-5. Disponible sur: <https://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2011.00131.x>
112. Cramb SM, Mengersen KL, Baade PD. Developing the atlas of cancer in Queensland: methodological issues. *Int J Health Geogr* [Internet]. 2011 [cité 16 nov 2022];10(1):9. Disponible sur: <http://ij-healthgeographics.biomedcentral.com/articles/10.1186/1476-072X-10-9>
113. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter DJ. *The BUGS book: a practical introduction to Bayesian analysis*. Boca Raton: CRC Press, Taylor & Francis Group; 2013.
114. Wakefield JC, Best NG, Waller L. Bayesian approaches to disease mapping. In: Elliott P, Wakefield J, Best N, Briggs D, éditeurs. *Spatial Epidemiology* [Internet]. 1<sup>re</sup> éd. Oxford University PressOxford; 2001 [cité 16 nov 2022]. p. 104-27. Disponible sur: <https://academic.oup.com/book/8189/chapter/153721813>
115. Mollié A. Bayesian mapping of Hodgkin's disease in France. In: *Spatial epidemiology: methods and applications*. Reprint. Oxford: Oxford Univ. Press; 2006. p. 267-85.
116. Lawson A, Browne WJ, Vidal Rodeiro CL. *Disease mapping with WinBUGS and MLwiN*. Chichester, West Sussex, England ; Hoboken, NJ: J. Wiley; 2003. 277 p. (Statistics in practice).

## Annexe 1 : Regroupements d'Iris peu ou pas peuplés

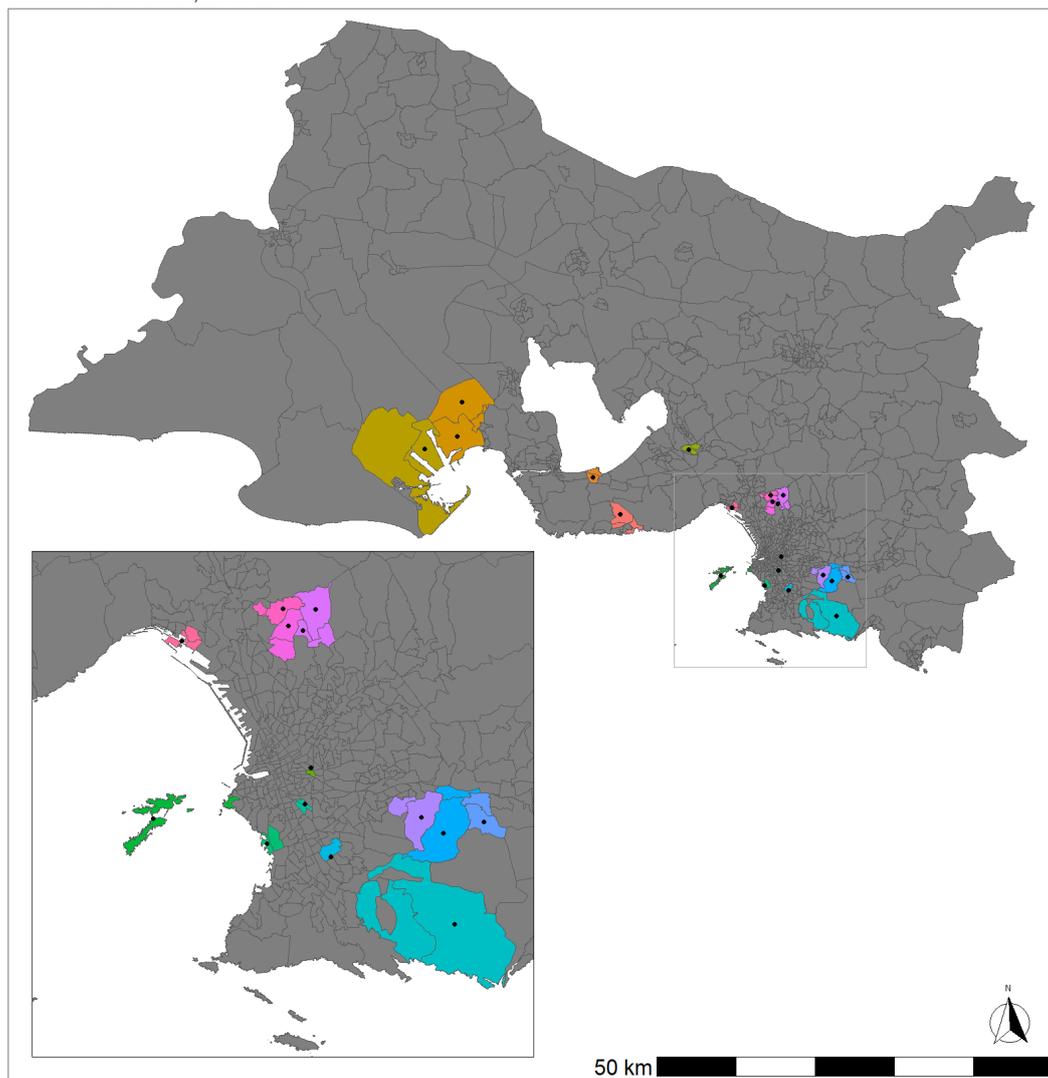
Le seuil retenu pour la définition des Iris peu peuplés était d'une population masculine ou féminine de moins de 25 personnes. L'Iris voisin le plus proche était identifié par la distance euclidienne minimale entre les barycentres géographiques des Iris. L'Iris « Les îles », regroupant les différentes îles au large de Marseille était groupé à l'Iris le plus proche (« Malmousque-Pêcheurs ») afin de pouvoir l'intégrer aux analyses utilisant une définition de voisinage se basant sur la contiguïté.

Au total, 20 Iris ont été fusionnés avec d'autres, dont 14 dans la ville de Marseille. Le total d'Iris utilisés dans l'analyse a donc diminué de 840 à 820. Les groupes ainsi fusionnés étaient inclus dans l'analyse sous le nom de l'Iris de regroupement.

La carte ci-dessous présente le résultat des regroupements effectués : les Iris fusionnés sont représentés par une même couleur et les points noirs indiquent les Iris peu ou pas peuplés. La liste complète des Iris concernés et leur Iris de regroupement est présentée à la suite.

**Carte des regroupements des Iris peu ou pas peuplés (< 25 hommes ou femmes) avec leurs plus proches voisins (les points noirs représentent les Iris peu ou pas peuplés, les couleurs permettent d'identifier le contour des nouvelles unités spatiales définies)**

Bouches-du-Rhône, 2013-2018



Point noir = IRIS peu ou pas peuplé (moins de 25 hommes ou femmes)  
Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - juin 2022

Liste des Iris peu ou pas peuplés (< 25 hommes ou femmes) et leur Iris de regroupement selon la nomenclature de l'Insee

IRIS peu ou pas peuplé			IRIS de regroupement		
Code IRIS	Code Insee de la commune	Nom de l'IRIS	Code IRIS	Code Insee de la commune	Nom de l'IRIS
130390104	13039	Brûle-Tabac-Roque-Galejon	130780101	13078	Malebarge
130390103	13039	Audience-Fenouillère-Cavaou-Clapet	130390105	13039	Crottes-Courbedonne-Carabins-Plaine Ronde
130390102	13039	Ventillon-Fossette-Grand Pati	130390105	13039	Crottes-Courbedonne-Carabins-Plaine Ronde
130260106	13026	Le Clos	130260105	13026	La Béarde
130210104	13021	Nord Romanon	130210103	13021	Est Reganat d'Estral
131170206	13117	L'Anjoly	131170209	13117	La Tuilière
132070301	13207	Les Îles	132070202	13207	Malmousque-Pêcheurs
132160402	13216	Bassins Mirabeau	132160401	13216	Mourepiane
132080502	13208	Le Palm Beach	132080501	13208	La Plage
132100202	13210	Parc du Prado	132080702	13208	Roger Renzo
132050201	13205	George-Verdun	132050206	13205	Barry-Camas
132140603	13214	Roches Claires	132140601	13214	Fontainieu
132150104	13215	Vallon Giraudy	132150201	13215	La Savine
132140703	13214	Carrière de Sainte-Marthe	132140401	13214	La Batarelle
132140405	13214	Vallon Dol	132140401	13214	La Batarelle
132090409	13209	E.D.F. Viton	132090809	13209	Les Hôpitaux-C.N.R.S.
132090902	13209	Calanque de Port Pin	132090603	13209	Calanque de Morgiou Partie Est
132110706	13211	Les Collines de Saint-Marcel	132111004	13211	Les Hauts de Saint-Jacques
132110202	13211	Les Collines de la Barasse	132110201	13211	La Barasse
132110502	13211	Les Collines de la Millière	132110501	13211	La Millière

## Annexe 2 : Établissements organisés par réunion de concertation pluridisciplinaire (RCP), par secteur géographique et par statut

### Liste des établissements et RCP d'urologie/oncologie par secteurs géographiques

Établissement	RCP	Statut	Secteurs
<b>Département des Bouches-du-Rhône</b>			
APHM - Hôpital Nord - Hôpital de la Conception	RCP Urologie Nord RCP Urologie Conception	Public	Marseille
Institut Paoli Calmettes	RCP Urologie IPC	CLCC	
Hôpital Saint-Joseph	RCP Urologie Prado-Louvain	Privé	
Clinique Chanteclerc	RCP Urologie Prado-Louvain	Privé	
Hôpital Européen	RCP Urologie Hôpital Européen	Privé	
Clinique Bouchard	RCP Urologie Hôpital Européen RCP Urologie Beauregard-Vert Coteau	Privé	
Clinique Bonneveine	RCP Urologie Prado-Louvain RCP Urologie Beauregard-Vert Coteau	Privé	
Hôpital privé de la Résidence du Parc	RCP Urologie IPC	Privé	
Clinique Vert Coteau	RCP Urologie Beauregard-Vert Coteau	Privé	
Hôpital privé Beauregard		Privé	
Hôpital privé Clairval	RCP Oncologie Clairval	Privé	
Centre hospitalier du Pays d'Aix	RCP Urologie du Pays d'Aix	Public	Aix en Provence
Clinique Axium	RCP Urologie Axium	Privé	
Polyclinique du Parc Rambot (PPR)	RCP de Pneumologie-Urologie PPR	Privé	
Centre hospitalier de Salon de Provence	RCP Oncologie CH Salon Provence RCP Polyvalente CH Salon Provence	Public	Salon de Provence
Clinique Vignoli	RCP Urologie Vignoli	Privé	
Centre hospitalier de Martigues	RCP Hôpital Nord	Public	Martigues
Clinique générale de Marignane	RCP Oncologie Roleb	Privé	
Clinique de l'Étang de l'Olivier	RCP Oncologie polyvalente HP Istres RCP Oncologie Roleb	Privé	
Clinique chirurgicale de Martigues	RCP Clinique Martigues	Privé	
Centre hospitalier Joseph Imbert	RCP Urologie Public Nîmes	Privé	Arles
Clinique Jeanne d'Arc	RCP Urologie Vaucluse	Privé	
Centre hospitalier Edmond Garcin	RCP Urologie Conception	Privé	Aubagne – La Ciotat
Clinique de La Ciotat	RCP Urologie Hôpital Européen RCP Urologie Beauregard-Vert Coteau	Privé	
Centre hospitalier de La Ciotat	RCP Urologie Hôpital Européen	Public	
Clinique de Casamance	RCP Urologie Prado-Louvain	Privé	
<b>Hors département des Bouches-du-Rhône</b>			
Polyclinique Kennedy	RCP Urologie Privé Nîmes	Privé	Nîmes
Centre hospitalier régional de Nîmes	RCP Urologie Public Nîmes	Public	
Centre hospitalier d'Avignon	RCP Oncologie	Public	Avignon
Institut Sainte Catherine	RCP Oncologie	Privé	
Clinique de Rhône-Durance	RCP Urologie Vaucluse	Privé	

## Liste des établissements et RCP d'hématologie/oncologie par secteurs géographiques

Établissement	RCP	Statut	Secteurs
<b>Département des Bouches-du-Rhône</b>			
AP-HM : Hôpital de la Conception	RCP Hématologie Conception	Public	Marseille
Institut Paoli-Calmettes	RCP Hématologie	Privé	
Clinique Vert Coteau	RCP Hématologie	Privé	
Hôpital privé Clairval	RCP Oncologie	Privé	Aix-en-Provence
Centre hospitalier du Pays d'Aix	RCP Hématologie	Public	
Hôpital privé de Provence	RCP Oncologie	Privé	
Centre hospitalier de Salon de Provence	RCP Oncologie /RCP Polyvalente	Public	Salon de Provence
Centre hospitalier de Martigues	RCP Oncologie Martigues	Public	Martigues
Centre hospitalier Joseph Imbert	RCP Hématologie	Public	Arles
Centre hospitalier Edmond Garcin	RCP Hématologie Conception	Public	Aubagne - La Ciotat
<b>Hors département des Bouches-du-Rhône</b>			
Centre hospitalier Avignon	RCP Hématologie Vaucluse	Public	Avignon

## Annexe 3 : Modèle de lissage spatial BYM

### 1. Présentation du modèle et de la méthode de calcul

Pour une unité géographique donnée, on définit  $O_i$  comme le nombre de cas (de cancer) observé et  $E_i$  le nombre de cas attendus. Dans notre étude, le nombre de cas attendus est calculé en effectuant une standardisation sur le nombre de personne-année par classe d'âge en utilisant comme référence les taux d'incidence des différents cancers par sexe pour la France métropolitaine sur la période d'étude.  $E_i$  correspond donc au nombre de cas attendus sous l'hypothèse nulle d'un taux d'incidence standardisé (SIR) égal à 1.

Le modèle hiérarchique bayésien proposé par Besag, York et Mollié en 1991 (24) combine un terme d'hétérogénéité non structurée (non spatial) et un terme d'hétérogénéité structurée indépendants l'un de l'autre. Il permet donc d'obtenir un lissage mixte : un compromis entre lissage global (obtenu par un modèle Poisson-log-normal par exemple) et lissage local.

Le modèle est caractérisé par (23) :

- un premier niveau (variabilité locale) : la vraisemblance qui modélise la structure des observations. Le nombre observé de cas de cancer  $O_i$  suit une distribution de Poisson de moyenne  $E_i\theta_i$  :

$$(O_i|\theta_i) \sim \text{Poisson}(E_i\theta_i)$$

Avec :

- $\theta_i$  le vrai risque relatif dans l'unité géographique
- L'estimateur  $\hat{\theta}_i = \frac{O_i}{E_i}$  correspondant au SIR

- un deuxième niveau (structure inter-zones) : la loi *a priori* des risques relatifs qui résume une information globale sur la similarité des risques  $\theta_i$ , sur leur moyenne et leur variabilité. Ce niveau permet d'introduire la dépendance spatiale :

$$\log(\theta_i) = \beta_0 + U_i + V_i + \sum_{k=1}^N \beta_k X_{ki}$$

Avec :

- $\beta_0$  Terme constant sur tout le territoire (moyenne)
- $U_i$  Effet aléatoire qui contrôle la variabilité des SIR dans sa composante non spatiale globale. Une valeur faible de  $\sigma_u^2$  indique des risques similaires entre unités spatiales (faible hétérogénéité spatiale).

$$U_i \sim \text{Normale}(0, \sigma_u^2)$$

- $V_i$  Effet aléatoire qui contrôle la variabilité des SIR dans sa composante spatiale. Cet effet suppose que les unités spatiales proches ont tendance à avoir un risque similaire. On utilise le modèle gaussien autorégressif intrinsèque (ICAR) pour prendre en compte la structure spatiale des données :

$$(V_i|V_j = v_j, j \neq i) \sim \text{Normale}\left(\frac{\sum_{j \neq i} w_{ij} v_j}{\sum_{j \neq i} w_{ij}}, \frac{\sigma_v^2}{\sum_{j \neq i} w_{ij}}\right)$$

Avec :

- $w_{ij}$  qui définit la notion de voisinage entre les unités spatiales, avec, dans le cas de l'utilisation de la contiguïté comme notion de voisinage,  $w_{ij} = 1$  si  $i$  et  $j$  partagent une frontière et 0 sinon
- $\sigma_v^2$  qui permet de contrôler la variabilité du risque dans sa composante spatiale

En d'autres termes, le modèle BYM suppose que la distribution conditionnelle de l'effet  $V_i$  dans l'unité spatiale  $i$  suit une loi normale centrée sur la moyenne des effets de ses unités voisines et de variance inversement proportionnelle au nombre de voisins. Une valeur faible de  $\sigma_v^2$  indique des risques similaires entre unités spatiales voisines.

- $\beta_k$  et  $X_{ki}$  respectivement coefficients linéaires et variables d'ajustement

Le résultat attendu est la distribution *a posteriori* des  $\log(\theta_i)$  (ainsi que celles des différents paramètres du modèle). La distribution *a posteriori* est le produit de la distribution *a priori* et de la fonction de vraisemblance. Si les données sont informatives, la vraisemblance dominera la valeur estimée des SIR  $\theta_i$  ; dans le cas contraire, l'information apportée par la loi *a priori* aura un poids plus important. Les modèles hiérarchiques bayésiens permettent d'intégrer ces deux types d'information. L'estimation des paramètres de ce modèle fait appel à des intégrales qui ne sont pas calculables par des méthodes analytiques. Il est nécessaire de faire appel à des méthodes de simulation (algorithmes stochastiques de Monte Carlo par Chaînes de Markov - McMC). Le package CARBayes permet dorénavant de réaliser l'intégralité de l'analyse directement sous R (33). La fonction S.CARbym a été utilisée pour effectuer l'ensemble des calculs.

Une fois la convergence du modèle atteinte, les estimations *a posteriori* des SIR  $\hat{\theta}_i$  sont obtenues en identifiant la médiane des valeurs de  $N$  simulations. En pratique, la sélection de ces  $N$  simulations est faite après avoir écarté les  $K$  premières simulations avant convergence (« burn-in period »). Une étape supplémentaire de sous-sélection de simulations avec un pas régulier (« thinning ») dans le but de réduire l'autocorrélation de simulations successives par l'algorithme McMC est souvent effectuée (111). Néanmoins, il a été pointé que cette pratique était « souvent non nécessaire et toujours inefficace » (111) à cause de la perte d'information ( $N/p$  simulations utiles au lieu de  $N$ ) qui n'est pas toujours compensée par l'amélioration de l'indépendance visée des simulations pour un nombre de simulations donné. Nous avons donc fait le choix de ne pas y avoir recours.

Les estimations *a posteriori* des SIR correspondent à l'exponentielle des médianes des distributions *a posteriori*. La même procédure est utilisée pour la détermination des risques relatifs correspondant aux effets des différentes covariables.

L'intervalle de crédibilité à 95 % des estimations, similaire à l'intervalle de confiance dans un contexte fréquentiste, est calculé en identifiant les percentiles à 2,5 et 97,5 de la distribution des  $N$  valeurs. De la même façon, une transformation exponentielle est utilisée pour déterminer l'intervalle de crédibilité à 95 % des SIR et des risques relatifs associés aux covariables.

## 2. Sélection du modèle et des paramètres

Les estimations *a posteriori* du modèle vont dépendre de plusieurs éléments propres au modèle lui-même, aux distributions *a priori* utilisées et au paramétrage utilisé pour le calcul des itérations. Colonna et Sauleau ont étudié la question de l'interprétation et du choix d'un modèle dans la description spatiale de cancers et offrent des orientations sur le processus de décision notamment dans le cadre de l'élaboration d'un atlas de cancer (28). Colonna a également étudié l'influence de différentes distributions *a priori* des variances des composantes aléatoires spatiales et non-spatiales (27). Le travail méthodologique effectué autour de l'élaboration de l'atlas des cancers du Queensland qui a ensuite été étendu à l'Australie entière (<https://atlas.cancer.org.au/>) fournit un exemple de mise en œuvre de ce processus de décision en apportant d'autres éléments (112).

Le critère principal utilisé pour comparer des modèles est le *Deviance information criterion* (DIC) qui est utilisé de la même façon que l'AIC (29). En suivant les recommandations de Lunn *et al.* (113) utilisées par Colonna et Sauleau (28), nous avons considéré le rejet d'un modèle pour un écart de 10 ou plus. Un écart de 5 à 10 était considéré comme une différence modérée nécessitant de regarder d'autres critères de choix et un écart de moins de 5 comme ne permettant pas de choisir.

Pour l'ensemble des cancers étudiés et pour les deux sexes, le modèle incluant l'ensemble des covariables était préféré s'il présentait un DIC ne dépassant pas de 10 le DIC du meilleur modèle. Ce choix reflétait l'objectif épidémiologique de pouvoir discuter des effets des différentes covariables et notamment des facteurs de risque connus ou supposés des différents cancers.

Afin d'évaluer la qualité des estimations du modèle, il est important d'évaluer la convergence des différents paramètres (113), et notamment celle des variances des composantes aléatoires. Elle peut s'estimer par le calcul de la taille d'échantillon effective (ou son pendant, l'erreur de Monte Carlo), qui prend en compte le nombre d'itérations (hors « burn-in ») et l'autocorrélation des estimations successives des itérations (113). La pratique retient un seuil d'une taille d'échantillon effective de 400 itérations (28). La convergence a été également vérifiée visuellement en vérifiant la ressemblance avec un « peigne » droit du tracé des estimations (113).

Plusieurs auteurs ont étudié l'influence des distributions *a priori* des variances des composantes spatiales et non-spatiales sur les estimations (27,114,115) et une analyse de sensibilité des estimations *a posteriori* à ces distributions a été également effectuée dans le cadre du développement de l'atlas de cancer du Queensland (113). En suivant l'exemple de Colonna (27), nous avons évalué l'influence de plusieurs distributions modélisant l'inverse des variances des hyper-paramètres proposées par plusieurs auteurs : Gamma(0.01,0.01) proposé par Wakefield *et al.* (115), Gamma(0.5,0.0005) utilisée par Lawson *et al.* (116) et les distributions empiriques dépendantes de la distribution des SIR observés introduites par Mollié (115).

Enfin, dans l'optique d'apporter de l'information pour choisir entre l'utilisation d'un modèle BYM ou d'un modèle de Poisson (sans composante spatiale), Colonna et Sauleau proposent de comparer la variance empirique de la composante spatiale et non spatiale afin d'évaluer si une des composantes apporte une part plus importante d'information (28). Dans le cas d'une prépondérance importante de la composante non spatiale, la nécessité de l'utilisation d'un modèle BYM peut être discutée.

### 3. Sélection du modèle et des paramètres

Après plusieurs essais, une période de « burn-in » de 5 000 itérations et 50 000 itérations additionnelles ont été jugées convenables. Ce nombre est un bon compromis entre le temps de calcul nécessaire pour l'ensemble des modèles et une bonne convergence des estimations. Aucun « thinning » n'a été effectué (paramètre fixé à 1) pour les raisons avancées dans la présentation de la méthode (111).

Pour les LAM chez les deux sexes, ainsi que pour le cancer du rein et de la vessie chez les femmes, le DIC le plus faible était retrouvé pour une distribution *a priori* Gamma(0.5,0.0005). Chez les hommes pour le cancer du rein, les distributions *a priori* empiriques de Mollié permettaient d'obtenir le DIC le plus faible. Finalement, pour le cancer de la vessie chez les hommes, le meilleur DIC était obtenu avec une distribution Gamma(0.1,0.1). Globalement, pour chaque cancer et sexe, l'ensemble des modèles (distributions *a priori* ou facteurs d'ajustement différents) présentaient des différences de DIC de moins de 10, sauf pour le cancer de la vessie chez l'homme. Pour ces modèles, l'absence de la variable de densité de

population induisait une différence supérieure à 10 quelle que soit la distribution *a priori* choisie.

Concernant la convergence, la distribution Gamma(0.1,0.1) permettait d'obtenir des tailles d'échantillon effectif équivalentes ou meilleures, notamment pour les deux variances des composantes aléatoires. La distribution Gamma(0.5,0.0005) présentait constamment des convergences mauvaises de ces deux paramètres. Les distributions empiriques de Mollié présentaient une qualité de convergence entre-deux, variant selon le cancer et le sexe. La vérification visuelle (forme de « peigne » (113)) confirmait clairement ces performances.

L'étendue des estimations *a posteriori* peut être influencée par les distributions *a priori* des variances des composantes spatiales et non spatiales (27,28). C'est d'ailleurs en utilisant l'étendue des données observées que sont calculées les distributions empiriques selon la méthode de Mollié (115). Nous n'avons pas observé de phénomène important de réduction de l'étendue des estimations *a posteriori* entre des modèles de DIC comparables, mais les étendues (quantiles à 2,5 % et 97,5 %) étaient généralement plus faibles pour les distributions Gamma(0.5,0.0005) et plus élevées pour Gamma(0.1,0.1). Dans le cadre

Enfin, le rapport des estimations empiriques des variances des deux composantes présentait des valeurs extrêmes pour des modèles utilisant la distribution Gamma (0.5,0.0005). Les distributions empiriques présentaient des rapports intermédiaires et la distribution Gamma (0.1,0.1) présentait les rapports les moins étendus variant de 1 à 4 en faveur de la composante spatiale. Les différences de convergence des estimations décrites ci-dessus offrent une bonne explication à ces différences de comportement entre les différentes distributions *a priori*. Pour les modèles utilisant la distribution Gamma(0.1,0.1), les rapports supérieurs à un semblent démontrer l'apport d'information de la composante spatiale du modèle, quel que soit le scénario. Cette information semble confirmer le choix d'un modèle BYM au détriment d'une modélisation de Poisson (28).

Finalement, la cartographie des médianes *a posteriori* des SIR (« SIR lissés ») ont été comparées pour les modèles complets de chaque configuration. Aucune différence visuelle notable des résultats obtenus n'a été observée.

La recherche d'une harmonisation de la démarche dans le cadre d'une analyse simultanée de plusieurs cancers tend à écarter le choix de la distribution de Mollié (27). Avec une meilleure convergence des estimations des variances et une étendue des estimations *a posteriori* plus large, nous avons fait le choix de retenir la distribution Gamma(0.1,0.1). Les DIC des modèles contenant l'ensemble des facteurs d'ajustement avaient des différences variant de 0 à 7,2 pour les six cas de figure, ce qui restait dans l'étendue acceptable.

À l'échelle des Iris, l'étendue attendue des estimations *a posteriori* des SIR attendue est équivalente ou plus large qu'à l'échelle communale puisqu'un plus grand contraste est attendu à une échelle spatiale plus fine. Puisqu'elle correspond à une étendue des estimations *a posteriori* plus large (27), la distribution *a priori* Gamma(0.1,0.1) a été sélectionnée pour effectuer le lissage sans effectuer d'analyses de sensibilité utilisant les autres distributions. Le nombre d'itérations a dû être augmenté à 200 000 itérations après le « burn-in » (5 000) pour assurer une convergence correcte des estimations et notamment des variances des composantes aléatoires. Le temps de calcul et la mémoire machine devenaient donc également des facteurs limitants.

Seules les données accumulées de cancer de la vessie chez les hommes remplissaient le critère de plus de 80 % des Iris présentant au moins un cas. Le lissage des SIR à l'échelle des Iris n'a donc été effectué que pour ce scénario. La convergence était correcte pour l'ensemble des modèles en utilisant 200 000 itérations : tailles d'échantillon effectives de plus de 200 pour la variance de la composante non spatiale et au-delà de 400 pour les autres paramètres. La

vérification visuelle de la convergence montrait des comportements satisfaisants. Le DIC du modèle contenant l'ensemble des covariables déviait de 7.7 du meilleur DIC (modèle contenant uniquement le *proxy* du tabagisme), et il a donc été considéré comme acceptable de le retenir pour pouvoir évaluer les effets des différentes covariables. Pour ce modèle, le rapport des estimations *a posteriori* des variances des composantes aléatoires était de 3,4 en faveur de la composante spatiale, ce qui justifiait *a posteriori* l'utilisation d'un effet aléatoire spatialisé pour effectuer le lissage.

## Annexe 4 : Méthodes de détection de clusters

### 1. Méthodes globales de clustering

Test de Potthoff et Whittinghil (34):

Sous l'hypothèse nulle, le nombre de cas observés  $O_i$  suit une distribution de Poisson de moyenne  $E_i$  le nombre de cas attendus. La statistique de test est :

$$PW = \sum_i^N \frac{O_i(O_i - 1)}{E_i}$$

Sous l'hypothèse nulle, on a alors :

$$PW \sim \mathcal{N}\left(\frac{O_+(O_+ - 1)}{E_+}, 2(N - 1) \frac{O_+(O_+ - 1)}{E_+^2}\right)$$

Avec  $O_+$  la somme des cas observés,  $E_+$  la somme des cas attendus et  $N$  le nombre d'entités spatiales.

Statistique (I) de Moran (30):

$$I = \frac{N}{\sum_i \sum_j \omega_{ij}} \frac{\sum_i \sum_j \omega_{ij} (O_i - \bar{O})(O_j - \bar{O})}{\sum_i (O_i - \bar{O})^2}$$

Avec  $\omega_{ij}$  les poids utilisés pour définir la proximité entre deux entités  $i$  et  $j$ , et  $N$  le nombre d'entités. La proximité est définie dans l'étude par une matrice de contiguïté de type Queen (1 si une frontière est commune, 0 sinon).

Une p-valeur est estimée par une simulation de Monte-Carlo en réalisant  $P = 9\,999$  permutations puis en déterminant le rang de l'indice de Moran du jeu de données.

Empirical Bayes Index (35):

Le calcul de l'indice se présente sous une forme similaire que celui de l'indice de Moran :

$$I = \frac{N}{\sum_i \sum_j \omega_{ij}} \frac{\sum_i \sum_j \omega_{ij} z_i z_j}{\sum_i (z_i - \bar{z})^2} \quad \text{où} \quad z_i = \frac{SIR_i - SIR_+}{\sqrt{v_i}}$$

Avec :

$$v_i = \left( s_i^2 - \frac{N}{E_+} SIR_+ \right) + \frac{SIR_+}{E_i} \quad \text{et} \quad s_i^2 = \frac{\sum E_i (SIR_i - SIR_+)^2}{E_+}$$

Avec  $SIR_+$ , le rapport des cas observés totaux sur les cas attendus totaux ( $E_+$ ).

L'amélioration se situe dans la prise en compte de la taille différente des populations (ou du nombre de cas attendus), et donc également de la variabilité de la variance des SIR selon le nombre de cas attendus. Cela passe par la normalisation des SIR par leur écart type estimé.

Comme pour l'indice de Moran, une p-valeur est estimée par une simulation de Monte-Carlo en réalisant  $P = 9\,999$  permutations puis en déterminant le rang du jeu de données observées.

Statistique de Tango (T) (36) :

$$T = \sum_i \sum_j \omega_{ij} \left( \frac{O_i}{O_+} - \frac{E_i}{E_+} \right) \left( \frac{O_j}{O_+} - \frac{E_j}{E_+} \right)$$

Avec  $O_+$  la somme des cas observés,  $E_+$  la somme des cas attendus et  $\omega_{ij}$  les poids utilisés pour définir la proximité entre deux entités  $i$  et  $j$ .

L'inférence est réalisée par la méthode de Monte-Carlo en réalisant  $P = 9\,999$  simulations de données sous l'hypothèse nulle d'un risque constant (les proportions locales suivent alors une distribution multinomiale). Le rang est alors déterminé pour obtenir une p-valeur.

## 2. Statistique de scan

Dans la méthode de scan spatial développée par Kulldorff (38), les entités géographiques sont représentées par leurs centroïdes et la fenêtre de scan de base est un cercle, centré sur un de ces centroïdes et de rayon variable. L'ensemble des entités de la zone d'étude est donc réparti entre l'intérieur ou l'extérieure de la fenêtre. La taille (rayon) maximale de la fenêtre est définie par un pourcentage de population totale limitée à l'intérieure de la fenêtre (taille maximale de la fenêtre de scan). Les nombres de cas observés et attendus à l'intérieur et à l'extérieur de la fenêtre sont comparés pour déterminer si cette fenêtre permet de définir un cluster.

La statistique de scan correspondante est basée sur un rapport de vraisemblance calculée pour chaque cluster potentiel. Dans le cas d'une distribution de Poisson, elle est proportionnelle à :

$$S \propto \left( \frac{O_{int}}{E_{int}} \right)^{O_{int}} \left( \frac{O_{ext}}{E_{ext}} \right)^{O_{ext}}$$

Le cluster le plus probable (ou cluster principal) est celui qui maximise cette quantité.

Pour la détection de clusters secondaires, les valeurs maximales suivantes sont utilisées (approche hiérarchique). De plus, la sélection se base généralement sur un second critère qui est la non-superposition (pas d'entités spatiales communes) du cluster suivant avec les clusters déjà sélectionnés.

Sans distribution connue pour la statistique, une p-valeur est obtenue pour chaque cluster par une simulation de Monte-Carlo, en simulant  $P = 9\,999$  jeux de données sous l'hypothèse d'un risque constant sur le territoire avec une distribution de Poisson.

Afin d'améliorer la détection des clusters de formes non régulières (40–42), notamment en bordure de territoire, une fenêtre elliptique a été utilisée (43). Par rapport à la fenêtre circulaire de base, définie uniquement par son centre (centroïde d'une entité) et le pourcentage de population totale recouverte (définissant un rayon), la fenêtre elliptique est définie par deux paramètres géométriques supplémentaires : son angle d'orientation et son excentricité (rapport entre le plus grand et le plus petit rayon). Une pénalisation de la statistique de scan basée sur cette excentricité  $s$  proposée par Kulldorff (43) a été utilisée pour favoriser les clusters compacts devant des clusters plus étirés à valeur égale de statistique:

$$S' \propto S \times \left( \frac{4s}{(s+1)^2} \right)$$

Afin de fournir l'information la plus pertinente, la sélection du meilleur jeu de clusters a été effectué par l'utilisation du coefficient de Gini (44) et non pas par la sélection hiérarchique sans superposition seule. Le coefficient de Gini (46) est un résumé de l'information du comportement d'une courbe de Lorenz (47) qui décrit et mesure l'hétérogénéité d'une distribution, originellement utilisée pour décrire la répartition des richesses. Han *et al.* ont adapté ce coefficient à la sélection de clusters détectés par statistique de scan (44). Cette méthode se base sur le constat que la taille maximale de la fenêtre de scan influe sur le résultat : une valeur trop grande peut engendrer la détection d'un très grand cluster principal pouvant cacher des clusters moyens plus informatifs ; à l'inverse, une taille trop petite peut entraîner l'éclatement d'un cluster réel en plusieurs très petits clusters peu informatifs.

La méthode consiste en une répétition de l'analyse à plusieurs tailles maximales de clusters (ex : 5, 10, 20, 30, 50 %) jusqu'à la taille maximale de fenêtre de scan (ex : 50 %), puis une sélection du meilleur jeu de cluster au sens de l'indice de Gini. En définissant un critère de sélection objectif en amont de l'analyse et en effectuant l'ensemble des analyses en une fois, cette méthode permet d'éviter les écueils des tests multiples et d'une sélection arbitraire selon les résultats observés. Une analyse des performances de cette méthode (45) a montré une meilleure capacité à identifier des clusters de formes irrégulières (non compacts), notamment dans sa version utilisant une fenêtre de scan elliptique que les versions classiques de la méthode.

Pour adapter l'indice de Gini à cette problématique, les auteurs introduisent plusieurs considérations (44) (ici pour une recherche de clusters de surincidence) :

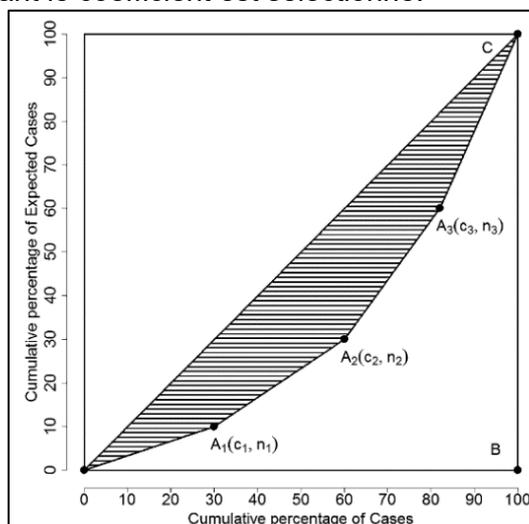
- 1) Les N clusters de chaque ensemble sont ordonnés par risques (ou SIR) décroissants
- 2) Chaque cluster se voit ensuite attribuer des coordonnées correspondant à la proportion cumulée de cas observés ( $c_i$ ) et à la proportion cumulée de cas attendus ( $n_i$ ) des clusters ( $A_1, \dots, A_i$ ).
- 3) L'origine (0,0) et le point (1,1) correspondent respectivement aux coordonnées ( $c_0, n_0$ ) et ( $c_{N+1}, n_{N+1}$ ).

Le coefficient de Gini s'écrit alors comme suit pour un jeu de clusters donné :

$$G = \sum_1^{N+1} (n_i c_{i-1} - n_{i-1} c_i)$$

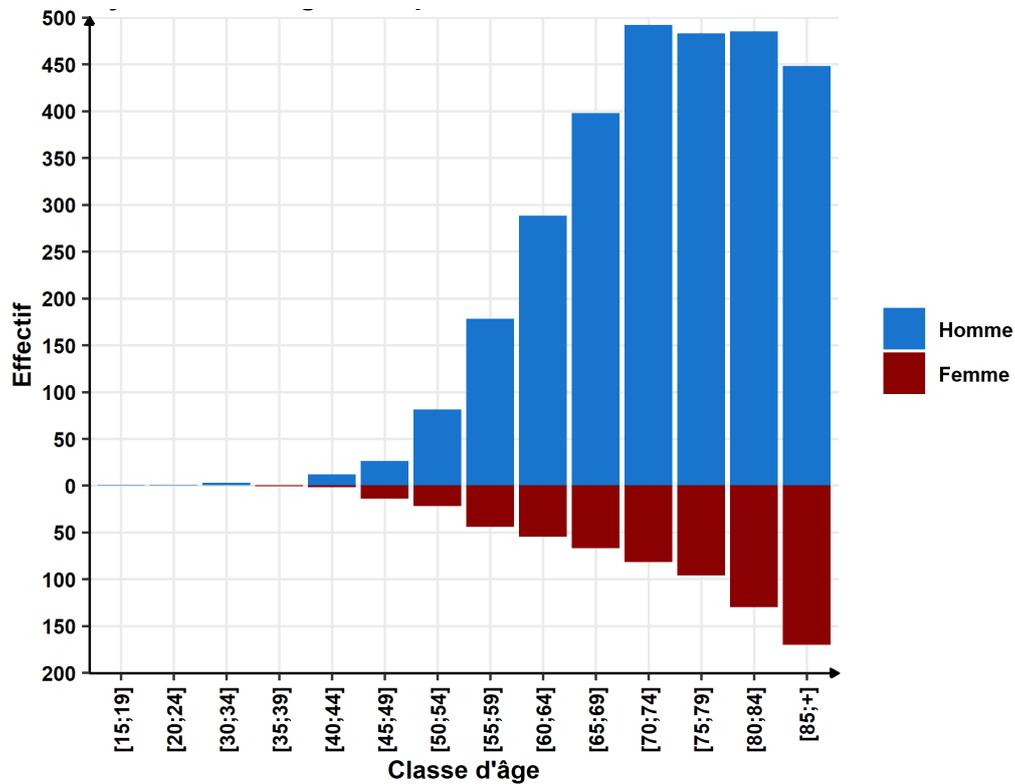
La figure ci-dessous extraite de (44) illustre la construction de la courbe de Lorenz à partir des points  $A_i$ . Le coefficient de Gini correspond à deux fois l'aire grisée, située entre la diagonale et cette courbe.

Le jeu de cluster maximisant le coefficient est sélectionné.

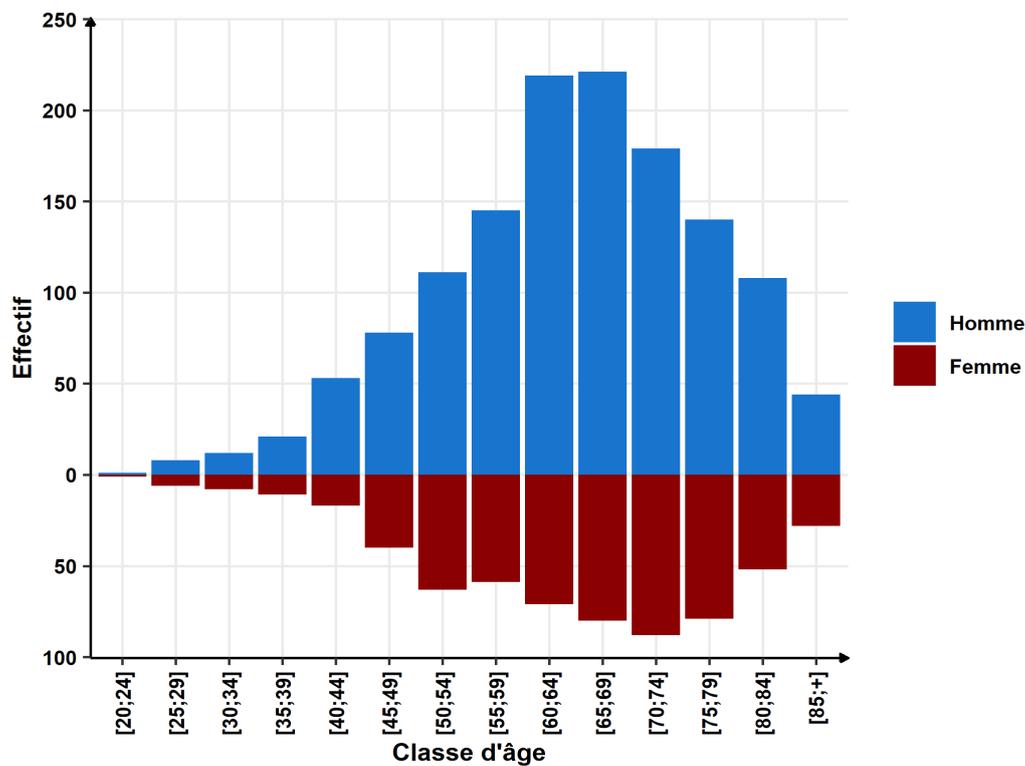


## Annexe 5 : Répartition des cas par sexe et classe d'âge

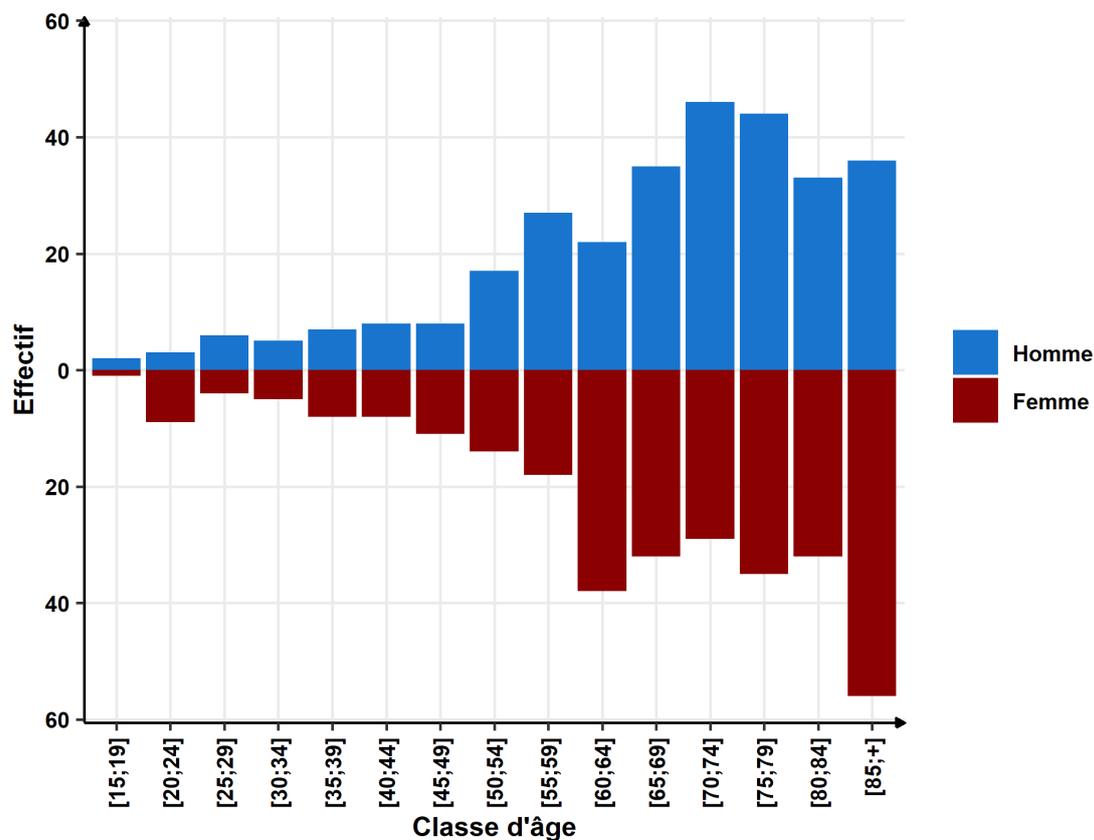
Répartition des cas de cancer urothéliaux par sexe et classe d'âge révolu au diagnostic, Bouches-du-Rhône, 2013-2018



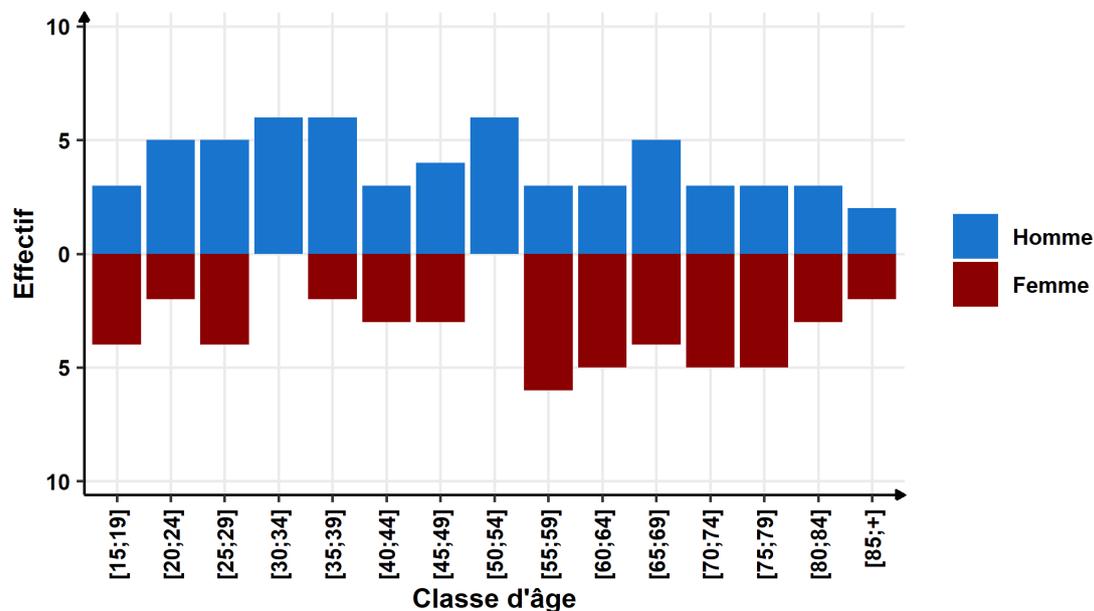
Répartition des cas de cancer du rein par sexe et classe d'âge révolu au diagnostic, Bouches-du-Rhône, 2013-2018



Répartition des cas de leucémie aiguë myéloïde (LAM) par sexe et classe d'âge révolu au diagnostic, Bouches-du-Rhône, 2013-2018



Répartition des cas de leucémie aiguë lymphoblastique (LAL) par sexe et classe d'âge révolu au diagnostic, Bouches-du-Rhône, 2013-2018



## Annexe 6 : Description des tumeurs urothéliales

### Répartition des tumeurs urothéliales par classe d'âge regroupée et par sexe, Bouches-du-Rhône, 2013-2018

	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 2 896	N = 683	N = 3 579	
<b>Classe d'âge</b>				<0,001
Moins de 65 ans	590 (20 %)	138 (20 %)	728 (20 %)	
65 - 84 ans	1 858 (64 %)	375 (55 %)	2 233 (62 %)	
85 ans et +	448 (15%)	170 (25%)	618 (17%)	

<sup>1</sup>Test du khi-deux d'indépendance

### Répartition des stades et grades des tumeurs urothéliales par sexe et classe d'âge, Bouches-du-Rhône, 2013-2018

	Moins de 65 ans, N = 728		65 - 84 ans, N = 2 233		85 ans et +, N = 618	
	Hommes, N = 590	Femmes, N = 138	Hommes, N = 1 858	Femmes, N = 375	Hommes, N = 448	Femmes, N = 170
<b>Stade tumoral</b>						
Stade pT1	334 (57 %)	64 (47 %)	1 089 (59 %)	187 (50%)	261 (58%)	70 (42%)
Stade > pT1	243 (41 %)	71 (52 %)	747 (40 %)	177 (47 %)	184 (41 %)	95 (57%)
Métastases prévalentes	11 (2 %)	1 (1 %)	18 (1 %)	10 (3 %)	3 (1 %)	3 (1 %)
Manquant	2	2	4	1	0	2
<b>Grade tumoral</b>						
Faible grade	14 (2 %)	4 (3 %)	27 (1 %)	14 (4 %)	6 (1 %)	1 (1 %)
Bas grade	107 (19 %)	27 (22 %)	318 (18 %)	54 (15 %)	67 (16 %)	18 (11 %)
Haut grade	443 (79 %)	92 (75 %)	1 441 (81 %)	288 (81%)	358 (83%)	139 (88%)
Manquant	26	15	72	19	17	12

## Annexe 7 : Description des tumeurs rénales

Répartition des cancers du parenchyme rénal par sexe, classification CIM10 et type de diagnostic, Bouches-du-Rhône, 2013-2018

CIM10 - type de diagnostic	Hommes	Femmes	Total
	N = 1 366	N = 625	N = 1 991
C64 - Biopsie uniquement	214 (16 %)	97 (16 %)	311 (16 %)
C64 - Intervention	1 126 (82 %)	506 (81 %)	1 632 (82 %)
D41 - Imagerie	26 (2 %)	22 (3 %)	48 (2 %)

Répartition des cancers du parenchyme rénal par classe d'âge regroupée et par sexe des cas confirmés par histologie (C64), Bouches-du-Rhône, 2013-2018

	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 1 340	N = 603	N = 1 943	
<b>Classe d'âge</b>				0,2
Moins de 65 ans	648 (48 %)	276 (46 %)	943 (48 %)	
65 - 84 ans	648 (48 %)	299 (50 %)	947 (49 %)	
85 ans et +	44 (4%)	28 (4%)	72 (3%)	

<sup>1</sup>Test du khi-deux d'indépendance

Répartition des résultats de l'histologie des cancers du parenchyme rénal par sexe des cas confirmés par histologie (C64), Bouches-du-Rhône, 2013-2018

	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 1 323	N = 595	N = 1 918	
<b>Histologie</b>				0,001
Cellules claires	958 (72 %)	442 (74 %)	1 400 (73 %)	
Tubulo-papillaires	252 (19 %)	77 (13 %)	329 (17 %)	
Chromophile	74 (6 %)	53 (9 %)	127 (7 %)	
Sarcomatoides	4 (0 %)	2 (0 %)	6 (0 %)	
Autre	35 (3 %)	21 (4 %)	56 (3 %)	
Manquant	17	8	25	

<sup>1</sup>Test du khi-deux d'indépendance; test exact de Fisher

## Annexe 8 : Description des leucémies aiguës

### Répartition des leucémies aiguës myéloïdes par classe d'âge regroupée et par sexe, Bouches-du-Rhône, 2013-2018

Patients LAM	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 299	N = 300	N = 599	
<b>Âge au diagnostic</b>				0,5
Médiane (Étendue)	71 (59 – 79)	71 (60 – 82)	71 (59 – 81)	
<b>Classe d'âge</b>				0,018
Moins de 65 ans	105 (35 %)	116 (39 %)	221 (37 %)	
65 - 84 ans	158 (53 %)	128 (43 %)	286 (48 %)	
85 ans et +	36 (12%)	56 (19%)	92 (15%)	

<sup>1</sup>Test de Wilcoxon-Mann-Whitney; test du khi-deux d'indépendance

### Répartition des leucémies aiguës lymphoblastiques par classe d'âge regroupée et par sexe, Bouches-du-Rhône, 2013-2018

Patients LAL	Hommes	Femmes	Total	p-valeur <sup>1</sup>
	N = 60	N = 48	N = 108	
<b>Âge au diagnostic</b>				0,077
Médiane (Étendue)	47 (32 – 66)	61 (40 – 72)	54 (33 – 70)	
<b>Classe d'âge</b>				0,4
Moins de 65 ans	44 (73 %)	29 (60 %)	73 (68 %)	
65 - 84 ans	14 (23 %)	17 (35 %)	31 (29 %)	
85 ans et +	2 (3,3%)	2 (4,2%)	4 (3,7%)	

<sup>1</sup>Test de Wilcoxon-Mann-Whitney; test exact de Fisher

### Comparaison des répartitions par classe d'âge regroupée des leucémies aiguës lymphoblastiques et myéloïdes, Bouches-du-Rhône, 2013-2018

Patients LAL ou LAM	LAM	LAL	Total	p-valeur <sup>1</sup>
	N = 599	N = 108	N = 707	
<b>Classe d'âge</b>				<0,001
moins de 65 ans	221 (37 %)	73 (68 %)	294 (42 %)	
65 - 84 ans	286 (48 %)	31 (29 %)	317 (45 %)	
85 ans et +	92 (15%)	4 (3,7%)	96 (14%)	

<sup>1</sup>Test du khi-deux d'indépendance

## Annexe 9 : Description histologique des leucémies aiguës

Type de leucémie	Code CIMO3	Libellé	Avant 2016 (N = 3 361)	À partir de 2016 (N = 3 961)	Total (N = 7 321)
LA	98013	Leucémie aiguë indifférenciée	1 (0,3 %)	6 (1,5 %)	7 (1,0 %)
LA MIXTE	98053	Leucémie aiguë biphénotypique	2 (0,6 %)	1 (0,3 %)	3 (0,4 %)
	98083	Leucémie aiguë de phénotype mixte, B/myéloïde	0 (0 %)	1 (0,3 %)	1 (0,1 %)
LA/LL	98113	Leucémie/lymphome lymphoblastique B	0 (0 %)	23 (5,8 %)	23 (3,1 %)
	98123	Leucémie/lymphome lymphoblastique B avec t(9;22)(q34.1;q11.2); BCR-ABL1	1 (0,3 %)	20 (5,1 %)	21 (2,9 %)
	98143	Leucémie/lymphome lymphoblastique B avec t(12;21)(p13.2;q22.1); ETV6-RUNX1	0 (0 %)	0 (0 %)	0 (0 %)
	98153	Leucémie/lymphome lymphoblastique B avec hyperdiploïdie	0 (0 %)	1 (0,3 %)	1 (0,1 %)
	98163	Leucémie/lymphome lymphoblastique B avec hypodiploïdie	0 (0 %)	1 (0,3 %)	1 (0,1 %)
	98363	Leucémie lymphoblastique à cellules précurseurs B (LAL B)	35 (10 %)	0 (0 %)	35 (4,8 %)
	98373	Leucémie/lymphome lymphoblastique T	13 (3,9 %)	7 (1,8 %)	20 (2,7 %)
LA/lymphome	96873	Lymphome de Burkitt	6 (1,8 %)	7 (1,8 %)	13 (1,8 %)
	98263	Leucémie de Burkitt	5 (1,5 %)	3 (0,8 %)	8 (1,1 %)
LAM	98613	Leucémie aiguë myéloïde	18 (5,4 %)	30 (7,6 %)	48 (6,6 %)
LAM ACR	98653	Leucémie aiguë myéloïde avec t(6;9)(p. 23;q34.1) DEK-NUP214	1 (0,3 %)	0 (0 %)	1 (0,1 %)
	98663	Leucémie aiguë promyélocytaire avec PML/RARA	20 (6,0 %)	15 (3,8 %)	35 (4,8 %)
	98693	Leucémie aiguë myéloïde avec inv(3)(q21;q26.2) ou t(3;3)(q21;q26.2); GATA2, MECOM(EVI1)	3 (0,9 %)	2 (0,5 %)	5 (0,7 %)
	98713	Leucémie aiguë myéloïde avec inv(16)(p13.1;q22) ou t(16;16)(p13.1;q22); CBFβ-MYH11	7 (2,1 %)	11 (2,8 %)	18 (2,5 %)
	98773	Leucémie aiguë myéloïde avec mutation NPM1	0 (0 %)	25 (6,3 %)	25 (3,4 %)
	98783	Leucémie aiguë myéloïde avec mutation biallélique de CEBPA	0 (0 %)	2 (0,5 %)	2 (0,3 %)
	98963	Leucémie aiguë myéloïde, t(8;21)(q22;q22.1) ; RUNX1-RUNX1T1	10 (3,0 %)	5 (1,3 %)	15 (2,0 %)
	98973	Leucémie aiguë myéloïde, t(9;11)(P21.3;23.3) KMT2A-MLLT3	1 (0,3 %)	6 (1,5 %)	7 (1,0 %)
	99113	Leucémie myéloïde aiguë (mégacaryoblastique) avec t(1;22)(p13.3;q13.1); RBM15-MKL1	1 (0,3 %)	0 (0 %)	1 (0,1 %)
	99123	Leucémie aiguë myéloïde BCR-ABL1	0 (0 %)	1 (0,3 %)	1 (0,1 %)
	LAM MRC	98953	Leucémie aiguë myéloïde secondaire à une myélodysplasie	47 (14 %)	51 (13 %)
LAM NOS	98403	Leucémie érythroïde (M6)	6 (1,8 %)	6 (1,5 %)	12 (1,6 %)
	98673	Leucémie aiguë myélo-monocytaire M4	26 (7,7 %)	14 (3,5 %)	40 (5,5 %)
	98703	Leucémie aiguë à basophile	0 (0 %)	1 (0,3 %)	1 (0,1 %)
	98723	Leucémie aiguë myéloïde peu différenciée LAM0	27 (8,0 %)	20 (5,1 %)	47 (6,4 %)
	98733	Leucémie aiguë myéloïde sans maturation LAM1	32 (9,5 %)	14 (3,5 %)	46 (6,3 %)
	98743	Leucémie aiguë myéloïde avec maturation LAM2	21 (6,2 %)	48 (12 %)	69 (9,4 %)
	98913	Leucémie aiguë monocytique et monoblastique (LAM5)	38 (11 %)	32 (8,1 %)	70 (9,6 %)
	99103	Leucémie aiguë mégacaryoblastique LAM7	1 (0,3 %)	2 (0,5 %)	3 (0,4 %)
	99313	Panmyélose aiguë avec myélofibrose	0 (0 %)	4 (1,0 %)	4 (0,5 %)
	Sarcome	99303	Sarcome granulocytaire	0 (0 %)	0 (0 %)
TR LAM	99203	Tumeur myéloïde secondaire à une thérapie	12 (3,6 %)	35 (8,8 %)	47 (6,4 %)
Tumeur blastique plasmocytoïde	97273	Tumeur blastique plasmocytoïde	2 (0,6 %)	2 (0,5 %)	4 (0,5 %)

## Annexe 10 : Résultats des modèles de lissage

### Cancers de la vessie

#### Valeurs des coefficients de régression et des effets associés des modèles de lissage spatiaux BYM des cancers de vessie chez l'homme à l'échelle des communes, Bouches-du-Rhône, 2013-2018

Cancer: Vessie - Sexe: Homme Échelle communale		Coefficients log-linéaires <sup>1</sup>		Effets sur les SIR <sup>2</sup>	
Variable d'ajustement	Médiane <sup>1</sup>	Intervalle de crédibilité à 95 % <sup>1</sup>	Effet <sup>2</sup>	Intervalle de crédibilité à 95 % <sup>2</sup>	
Proxy du tabagisme	0,78	[ 0,18 ; 1,37]	<b>2,17</b>	<b>[1,20 ; 3,95]</b>	
Indice de défavorisation sociale - Fdep	-0,04	[-0,08 ; 0,00]	0,96	[0,92 ; 1,00]	
Indice Insee: 2 - Densité intermédiaire	0,33	[ 0,13 ; 0,54]	<b>1,39</b>	<b>[1,14 ; 1,72]</b>	
Indice Insee: 1 - Densément peuplées	0,54	[ 0,32 ; 0,77]	<b>1,72</b>	<b>[1,37 ; 2,16]</b>	
Accessibilité potentielle localisée	0,03	[-0,03 ; 0,09]	1,03	[0,98 ; 1,09]	

DIC: 681,8

Log-likelihood: -304,1

<sup>1</sup> Estimations calculées à partir des 50 000 itérations de la procédure Monte Carlo par chaînes de Markov (McMC) du modèle CAR BYM ; <sup>2</sup> Valeurs calculées par transformation exponentielle

#### Valeurs des coefficients de régression et des effets associés des modèles de lissage spatiaux BYM des cancers de vessie chez l'homme à l'échelle des Iris, Bouches-du-Rhône, 2013-2018

Cancer: Vessie - Sexe: Homme Échelle IRIS		Coefficients log-linéaires <sup>1</sup>		Effets sur les SIR <sup>2</sup>	
Variable d'ajustement	Médiane <sup>1</sup>	Intervalle de crédibilité à 95 % <sup>1</sup>	Effet <sup>2</sup>	Intervalle de crédibilité à 95 % <sup>2</sup>	
Proxy du tabagisme	0,37	[-0,09 ; 0,83]	1,45	[0,92 ; 2,29]	
Indice de défavorisation sociale - Fdep	0,00	[-0,03 ; 0,03]	1,00	[0,97 ; 1,03]	
Dens. de pop. (hab./km <sup>2</sup> ) - Quartile: ]466 ; 3077]	-0,01	[-0,14 ; 0,11]	0,99	[0,87 ; 1,12]	
Dens. de pop. (hab./km <sup>2</sup> ) - Quartile: ]3077 ; 7815]	-0,05	[-0,20 ; 0,09]	0,95	[0,82 ; 1,09]	
Dens. de pop. (hab./km <sup>2</sup> ) - Quartile: ]7815 ; 80600]	0,01	[-0,17 ; 0,17]	1,01	[0,84 ; 1,19]	
Accessibilité potentielle localisée	-0,01	[-0,05 ; 0,04]	0,99	[0,95 ; 1,04]	

DIC: 3176,3

Log-likelihood: -1493,6

<sup>1</sup> Estimations calculées à partir des 200 000 itérations de la procédure Monte Carlo par chaînes de Markov (McMC) du modèle CAR BYM ; <sup>2</sup> Valeurs calculées par transformation exponentielle

#### Valeurs des coefficients de régression et des effets associés des modèles de lissage spatiaux BYM des cancers de vessie chez la femme à l'échelle des communes, Bouches-du-Rhône, 2013-2018

Cancer: Vessie - Sexe: Femme Échelle communale		Coefficients log-linéaires <sup>1</sup>		Effets sur les SIR <sup>2</sup>	
Variable d'ajustement	Médiane <sup>1</sup>	Intervalle de crédibilité à 95 % <sup>1</sup>	Effet <sup>2</sup>	Intervalle de crédibilité à 95 % <sup>2</sup>	
Proxy du tabagisme	0,76	[-0,58 ; 2,16]	2,13	[0,56 ; 8,64]	
Indice de défavorisation sociale - Fdep	0,00	[-0,06 ; 0,05]	1,00	[0,94 ; 1,06]	
Indice Insee: 2 - Densité intermédiaire	0,24	[-0,23 ; 0,68]	1,27	[0,79 ; 1,97]	
Indice Insee: 1 - Densément peuplées	0,45	[-0,04 ; 0,88]	1,57	[0,96 ; 2,42]	
Accessibilité potentielle localisée	0,00	[-0,10 ; 0,09]	1,00	[0,90 ; 1,10]	

DIC: 448,3

Log-likelihood: -208,7

<sup>1</sup> Estimations calculées à partir des 50 000 itérations de la procédure Monte Carlo par chaînes de Markov (McMC) du modèle CAR BYM ; <sup>2</sup> Valeurs calculées par transformation exponentielle

## Cancers du rein

### Valeurs des coefficients de régression et des effets associés des modèles de lissage spatiaux BYM des cancers du rein chez l'homme à l'échelle des communes, Bouches-du-Rhône, 2013-2018

Cancer: Rein - Sexe: Homme Échelle communale		Coefficients log-linéaires <sup>1</sup>		Effets sur les SIR <sup>2</sup>	
Variable d'ajustement	Médiane <sup>1</sup>	Intervalle de crédibilité à 95 % <sup>1</sup>	Effet <sup>2</sup>	Intervalle de crédibilité à 95 % <sup>2</sup>	
Proxy du tabagisme	0,27	[-0,41 ; 0,95]	1,31	[0,67 ; 2,58]	
Indice de défavorisation sociale - Fdep	-0,01	[-0,06 ; 0,04]	0,99	[0,94 ; 1,04]	
Indice Insee: 2 - Densité intermédiaire	0,21	[-0,02 ; 0,44]	1,23	[0,98 ; 1,55]	
Indice Insee: 1 - Densément peuplées	0,23	[-0,02 ; 0,48]	1,25	[0,98 ; 1,61]	
Accessibilité potentielle localisée	-0,06	[-0,12 ; 0,00]	0,94	[0,89 ; 1,00]	

DIC: 618,2

Log-likelihood: -283,9

<sup>1</sup> Estimations calculées à partir des 50 000 itérations de la procédure Monte Carlo par chaînes de Markov (McMC) du modèle CAR BYM ; <sup>2</sup> Valeurs calculées par transformation exponentielle

### Valeurs des coefficients de régression et des effets associés des modèles de lissage spatiaux BYM des cancers du rein chez la femme à l'échelle des communes, Bouches-du-Rhône, 2013-2018

Cancer: Rein - Sexe: Femme Échelle communale		Coefficients log-linéaires <sup>1</sup>		Effets sur les SIR <sup>2</sup>	
Variable d'ajustement	Médiane <sup>1</sup>	Intervalle de crédibilité à 95 % <sup>1</sup>	Effet <sup>2</sup>	Intervalle de crédibilité à 95 % <sup>2</sup>	
Proxy du tabagisme	-1,45	[-2,56 ; -0,32]	<b>0,23</b>	<b>[0,08 ; 0,73]</b>	
Indice de défavorisation sociale - Fdep	0,03	[-0,02 ; 0,08]	1,03	[0,98 ; 1,08]	
Indice Insee: 2 - Densité intermédiaire	-0,24	[-0,53 ; 0,07]	0,79	[0,59 ; 1,08]	
Indice Insee: 1 - Densément peuplées	-0,28	[-0,57 ; 0,06]	0,76	[0,57 ; 1,06]	
Accessibilité potentielle localisée	0,01	[-0,07 ; 0,10]	1,01	[0,93 ; 1,10]	

DIC: 503,1

Log-likelihood: -236,2

<sup>1</sup> Estimations calculées à partir des 50 000 itérations de la procédure Monte Carlo par chaînes de Markov (McMC) du modèle CAR BYM ; <sup>2</sup> Valeurs calculées par transformation exponentielle

## Leucémies aiguës myéloïdes

### Valeurs des coefficients de régression et des effets associés des modèles de lissage spatiaux BYM des leucémies aiguës myéloïdes chez l'homme à l'échelle des communes, Bouches-du-Rhône, 2013-2018

Cancer: Lam - Sexe: Homme Échelle communale		Coefficients log-linéaires <sup>1</sup>		Effets sur les SIR <sup>2</sup>	
Variable d'ajustement	Médiane <sup>1</sup>	Intervalle de crédibilité à 95 % <sup>1</sup>	Effet <sup>2</sup>	Intervalle de crédibilité à 95 % <sup>2</sup>	
Proxy du tabagisme	0,32	[-1,04 ; 1,69]	1,38	[0,35 ; 5,41]	
Indice de défavorisation sociale - Fdep	0,08	[-0,02 ; 0,18]	1,09	[0,98 ; 1,20]	
Indice Insee: 2 - Densité intermédiaire	-0,05	[-0,53 ; 0,47]	0,96	[0,59 ; 1,61]	
Indice Insee: 1 - Densément peuplées	-0,07	[-0,56 ; 0,42]	0,93	[0,57 ; 1,52]	
Accessibilité potentielle localisée	0,07	[-0,05 ; 0,19]	1,08	[0,95 ; 1,21]	

DIC: 367,3

Log-likelihood: -171,0

<sup>1</sup> Estimations calculées à partir des 50 000 itérations de la procédure Monte Carlo par chaînes de Markov (McMC) du modèle CAR BYM ; <sup>2</sup> Valeurs calculées par transformation exponentielle

### Valeurs des coefficients de régression et des effets associés des modèles de lissage spatiaux BYM des leucémies aiguës myéloïdes chez la femme à l'échelle des communes, Bouches-du-Rhône, 2013-2018

Cancer: Lam - Sexe: Femme Échelle communale		Coefficients log-linéaires <sup>1</sup>		Effets sur les SIR <sup>2</sup>	
Variable d'ajustement	Médiane <sup>1</sup>	Intervalle de crédibilité à 95 % <sup>1</sup>	Effet <sup>2</sup>	Intervalle de crédibilité à 95 % <sup>2</sup>	
Proxy du tabagisme	0,64	[-1,11 ; 2,34]	1,89	[0,33 ; 10,37]	
Indice de défavorisation sociale - Fdep	-0,02	[-0,10 ; 0,06]	0,98	[0,91 ; 1,06]	
Indice Insee: 2 - Densité intermédiaire	0,79	[0,20 ; 1,56]	<b>2,21</b>	<b>[1,22 ; 4,78]</b>	
Indice Insee: 1 - Densément peuplées	0,74	[0,15 ; 1,55]	<b>2,10</b>	<b>[1,16 ; 4,70]</b>	
Accessibilité potentielle localisée	0,06	[-0,06 ; 0,18]	1,06	[0,94 ; 1,20]	

DIC: 342,7

Log-likelihood: -159,9

<sup>1</sup> Estimations calculées à partir des 50 000 itérations de la procédure Monte Carlo par chaînes de Markov (McMC) du modèle CAR BYM ; <sup>2</sup> Valeurs calculées par transformation exponentielle

## Annexe 11 : Analyse de sensibilité des refus

Cette analyse de sensibilité prend en compte les informations disponibles sur les cas suivis par des urologues ayant refusé que les données de leurs patients soient saisies dans la base de données de l'Observatoire REVELA13.

Des informations concernant le sexe, l'âge au diagnostic, le type de cancer et la commune de résidence de 22 patients des urologues ayant refusé de participer au recueil ont pu être collectées par d'autres sources médicales intervenant dans leur prise en charge. Ces patients n'étaient pas enregistrés dans la base de données.

Parmi ces 22 patients, 7 étaient des cas de cancer du parenchyme rénal (6 hommes et 1 femme) et 15 étaient atteints de cancers de la vessie (11 hommes et 4 femmes).

Les cas de cancers du rein se répartissaient entre les communes de Cabannes (1H), Lamanon (1H), Mallemort (1F), Plan d'Orgon (2H) et Senas (2H).

Les cas de cancers de la vessie se répartissaient entre les communes de Martigues (5H/1F), Senas (1H), Chateauneuf-les-Martigues (1F), Gignac-la-Nerthe (1H), Istres (1F), le 10<sup>e</sup> arrondissement de Marseille (1H/1F), Port-de-Bouc (1H/1F) et St Mitre-les-Remparts (1H).

Les cas se répartissaient par classe d'âge quinquennale comme suit :

Classe d'âge	Cancer du rein (N = 7)		Cancer de la vessie (N = 15)	
	Hommes N = 6	Femmes N = 1	Hommes N = 11	Femmes N = 4
[45;49]	2			
[50;54]	1		1	
[55;59]	1			1
[60;64]	2			
[65;69]		1	1	
[70;74]			1	1
[75;79]			2	
[80;84]			3	2
[85;+]			3	

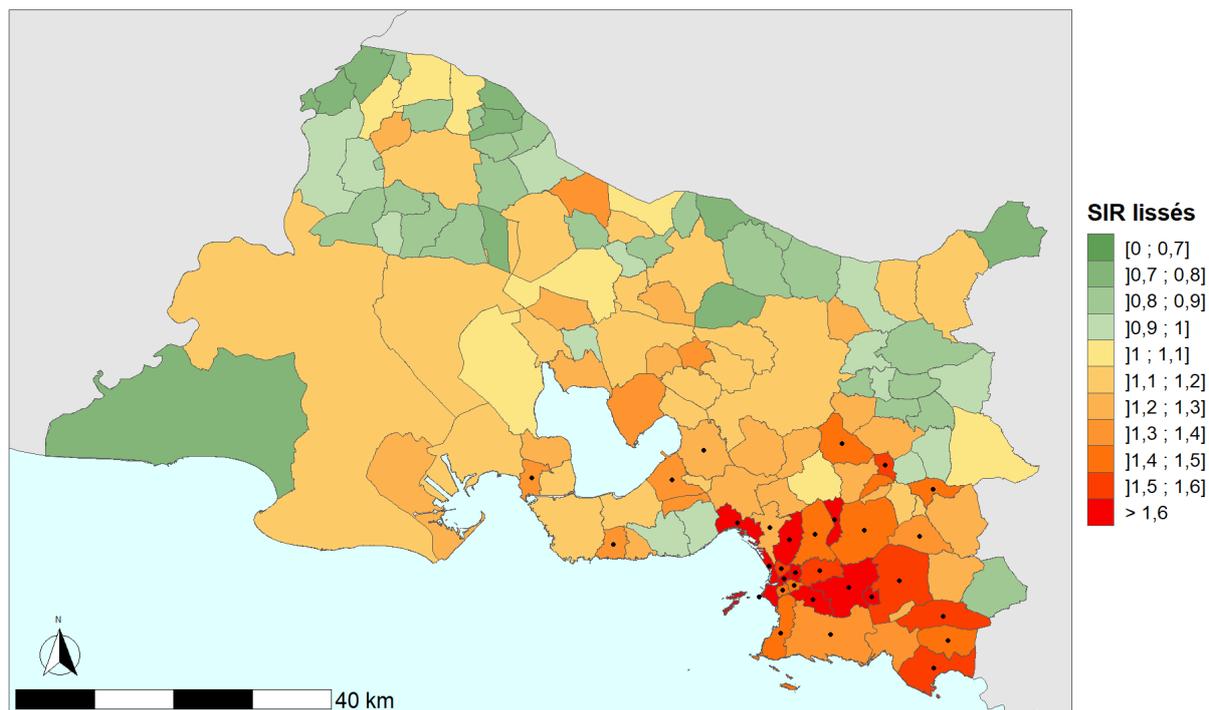
### *Cancers de la vessie : Répartition spatiale des risques*

#### **A) Résultats chez les hommes**

##### *Cartographie des risques lissés à l'échelle de la commune*

L'inclusion des cas de cancer de la vessie masculins dans l'analyse augmentait légèrement les valeurs des SIR lissés des communes où résidaient ces cas, notamment à Martigues avec 5 des 11 cas recensés. Néanmoins, cela n'engendrait aucun changement notable dans la distribution des surincidences significatives (figure A12-1).

**Figure (A12) - 1. Cartographie des SIR lissés de cancers de la vessie chez les hommes incluant les cas supplémentaires liés aux refus de participation au recueil selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

### *Identification de clusters à l'échelle de la commune*

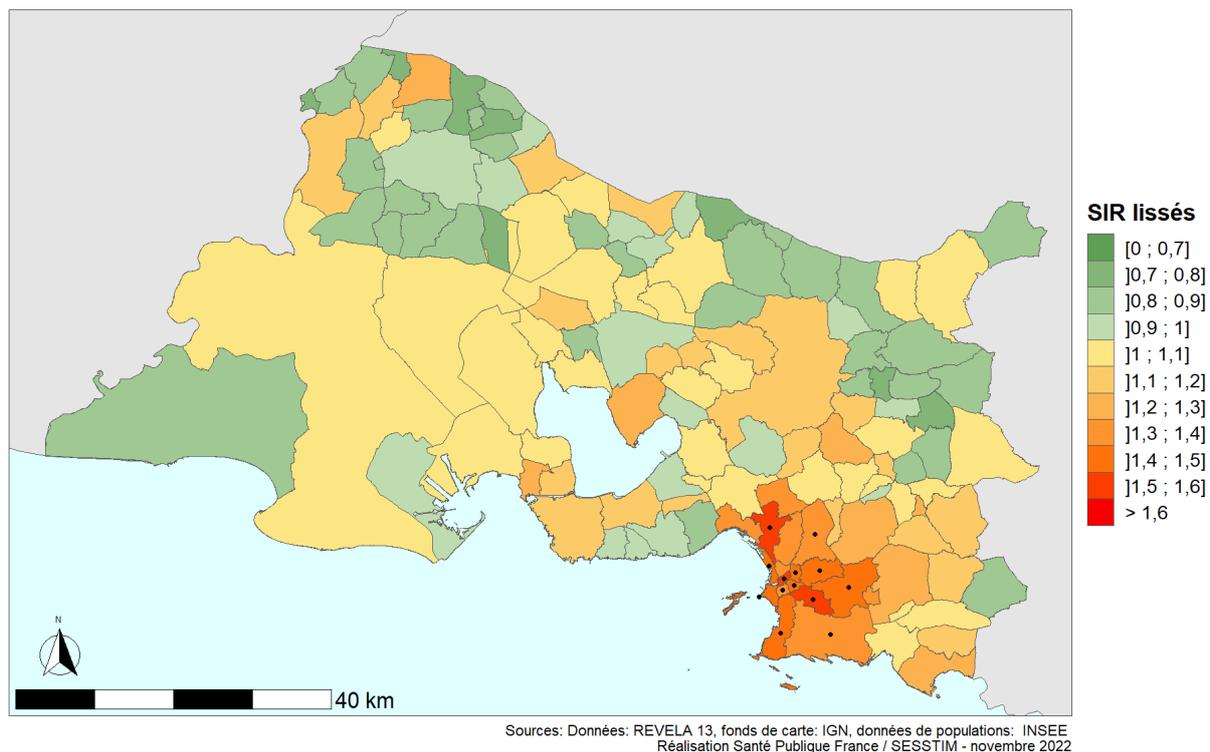
L'identification de clusters donnait des résultats identiques dans les analyses avec et sans l'inclusion de la commune de Marseille.

### **B) Résultats chez les femmes**

#### *Cartographie des risques lissés à l'échelle de la commune*

L'inclusion des cas de cancer de la vessie féminins dans l'analyse entraînait également une légère augmentation des valeurs de SIR lissées dans les communes concernées et leur voisinage, sans provoquer de changement dans la distribution des surincidences significatives (figure A12-2).

**Figure (A12) - 2. Cartographie des SIR lissés de cancers de la vessie chez les femmes incluant les cas supplémentaires liés aux refus de participation au recueil selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



### *Identification de clusters*

L'identification de clusters donnait des résultats identiques dans les analyses avec et sans l'inclusion de la commune de Marseille.

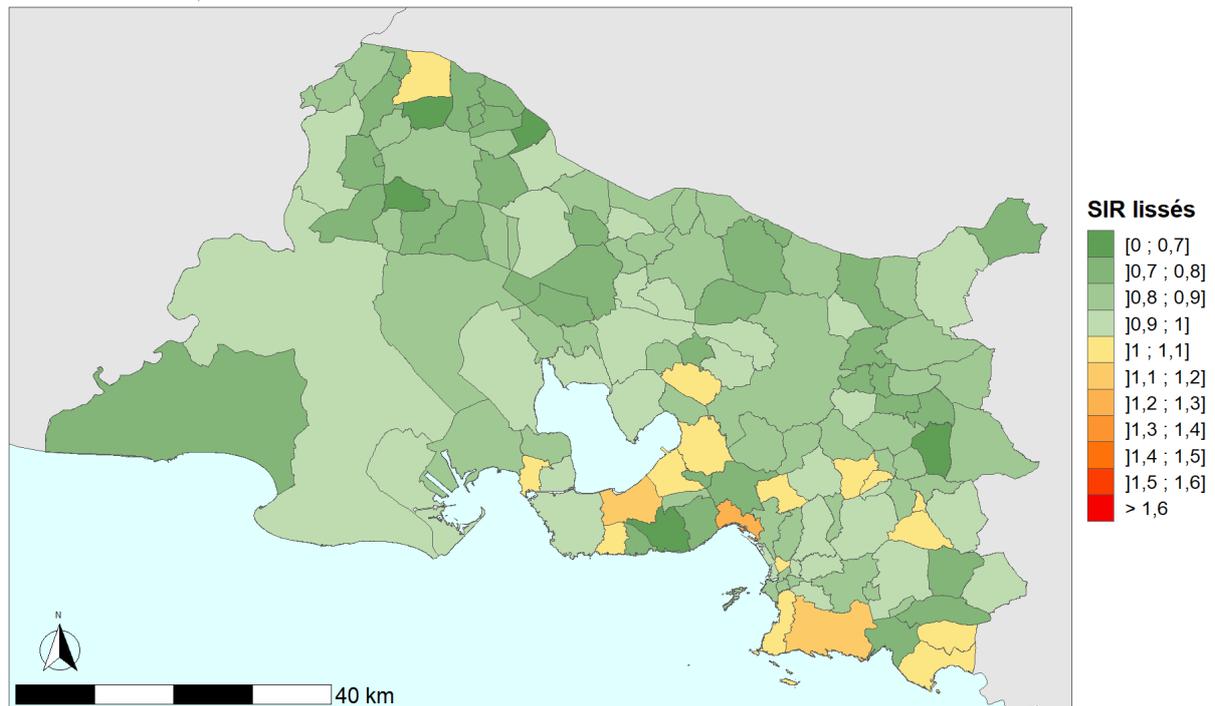
### *Cancers du rein : Répartition spatiale des risques*

#### **A) Résultats chez les hommes**

#### *Cartographie des risques lissés à l'échelle de la commune*

L'inclusion des 6 cas de cancer du rein masculins dans l'analyse ne modifiait pas les valeurs de SIR au-delà des variations aléatoires dues à la méthode de lissage non paramétrique. Aucune commune ne présentait de surincidence significative.

**Figure (A12) - 3. Cartographie des SIR lissés de cancers du rein chez les hommes incluant les cas supplémentaires liés aux refus de participation au recueil selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

#### *Identification de clusters à l'échelle de la commune*

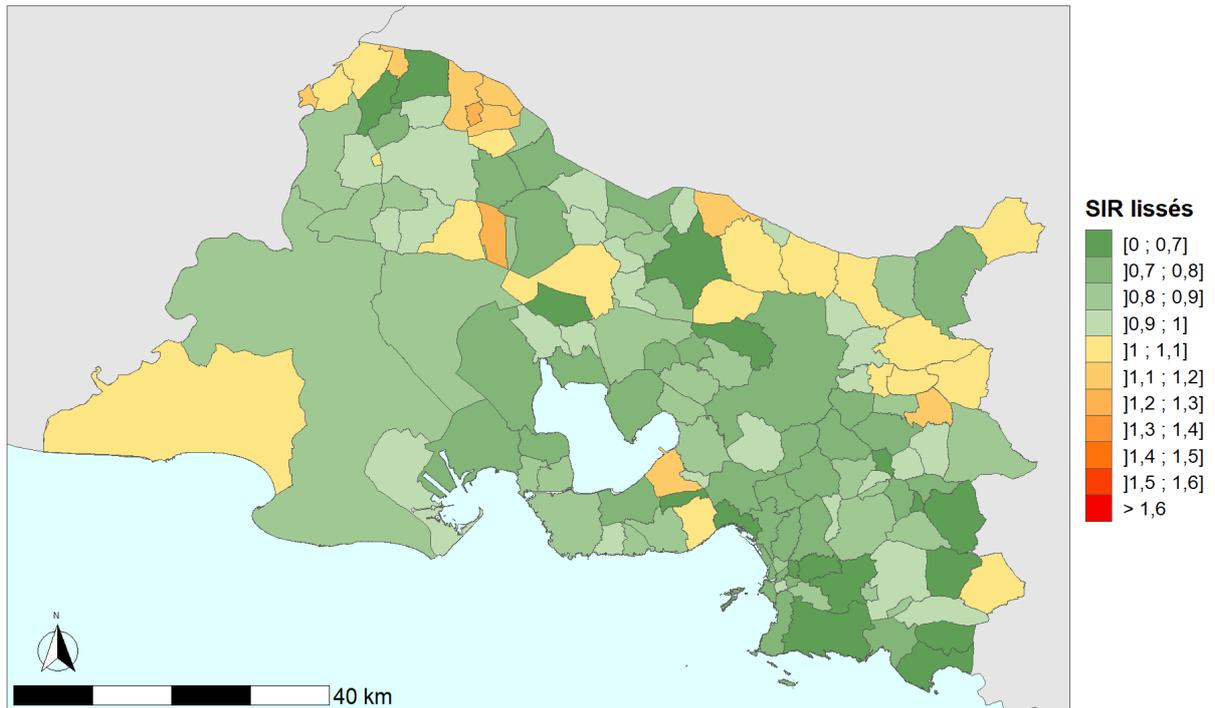
Comme dans l'analyse principale, aucun cluster significatif n'était détecté.

#### **B) Résultats chez les femmes**

##### *Cartographie des risques lissés à l'échelle de la commune*

L'inclusion du cas de cancer du rein féminin dans l'analyse ne modifiait pas les valeurs de SIR au-delà des variations aléatoires dues à la méthode de lissage non paramétrique. Aucune commune ne présentait de surincidence significative.

**Figure (A12) - 4. Cartographie des SIR lissés de cancers de vessie chez les femmes incluant les cas supplémentaires liés aux refus de participation au recueil selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

### *Identification de clusters*

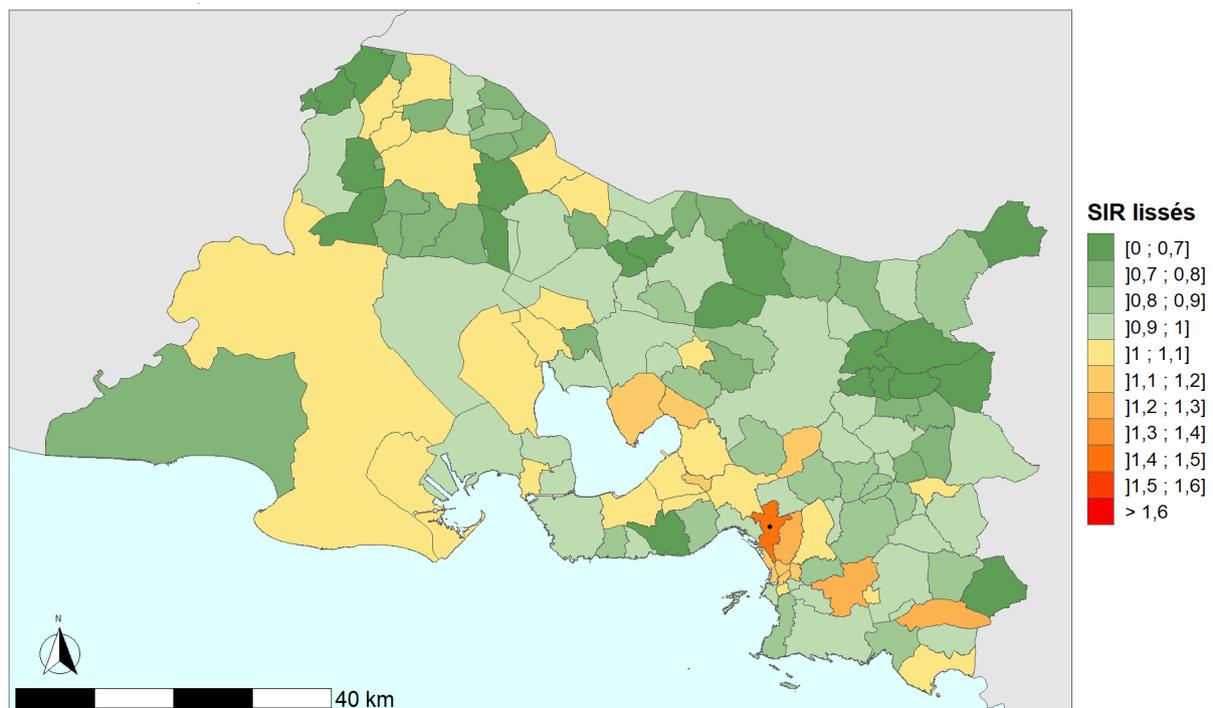
Comme dans l'analyse principale, aucun cluster significatif n'était détecté.

## Annexe 12 : Résultats de l'analyse de sensibilité des LAM

L'analyse conjointe des cas de LAM hommes et femmes visait à augmenter les effectifs de cas et de population, et donc à réduire l'étendue des intervalles de crédibilité à 95 % des SIR lissés, au détriment potentiellement de spécificités des répartitions spatiales par sexe qui pouvaient être estompées. En effet, on était en mesure d'attendre une amélioration de la détection de surincidences de risques significatifs principalement dans les zones où les deux sexes présentaient des valeurs de SIR élevées.

À l'échelle communale, seul le 15<sup>e</sup> arrondissement de Marseille présentait une surincidence significative, comme dans les résultats du lissage spatial chez les hommes. Par rapport à ce dernier, la valeur de SIR diminuait légèrement avec l'inclusion des cas et de la population féminines (de 1,51 CI95 % [1,00 ; 2,16] à 1,41 CI95 % [1,06 ; 1,85]).

**Figure (A13) - 1. Cartographie des SIR lissés de LAM tous sexes confondus selon la commune de résidence, Bouches-du-Rhône, 2013-2018**



Sources: Données: REVELA 13, fonds de carte: IGN, données de populations: INSEE  
Réalisation Santé Publique France / SESSTIM - novembre 2022

### Identification de clusters

Comme pour les analyses menées avec les deux sexes pris individuellement, aucun cluster n'avait été identifié.