

**VEILLE ET
SURVEILLANCE**

SEPTEMBRE 2018

ÉTUDES ET ENQUÊTES

OBSERVATOIRE DES CANCERS DU REIN,
DE LA VESSIE ET DES LEUCÉMIES
AIGUËS CHEZ LES ADULTES,
BOUCHES-DU-RHÔNE

Analyse de l'exhaustivité, 2013-2014

RÉGION

PROVENCE ALPES CÔTE D'AZUR

Résumé

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

L'observatoire Revela13 a été mis en œuvre en avril 2012 dans le département des Bouches-du-Rhône. L'enjeu était de tester des méthodes de recensement et de géolocalisation de ces cancers incidents et de réaliser des analyses spatiales pour identifier d'éventuels regroupements de cas sur le territoire concerné. Le recueil de données était basé sur le signalement par les praticiens des cas de cancers ciblés, complété par des extractions des fiches de réunion de concertation pluridisciplinaire (RCP) issues des centres de coordination en cancérologie. Après trois ans de recueil de données, une étude capture-recapture sur la période 2013-2014 a été menée pour évaluer l'exhaustivité de l'observatoire Revela13 en retenant comme sources complémentaires les données du Programme de médicalisation des systèmes d'informations (PMSI), et les données issues des laboratoires d'anatomocytopathologies (ACP). Une analyse à deux sources (Revela13 et PMSI) a été réalisée pour les leucémies aiguës et à trois sources pour les cancers urinaires.

Les deux sources de données complémentaires ont permis de récupérer 15 % (cancers urinaires) à 20 % (leucémies aiguës) de cas supplémentaires sur la période 2013-2014. Le nombre de cas estimé était de 232 [IC 95% : 225-239] pour les leucémies aiguës, de 1 247 [1 239-1 255] pour les cancers urothéliaux et de 600 [589-611] pour les cancers du rein. L'exhaustivité de l'observatoire sans stratification était respectivement de 76 % [74 %-79 %], de 84 % [83 %-85 %] et de 84 % [83 %-86 %]. Pour les trois cancers, les personnes les plus âgées avaient significativement moins de chance d'être captées par l'observatoire Revela13 que les plus jeunes. Pour les cancers urothéliaux, le stade et la zone de résidence étaient également à l'origine d'une hétérogénéité de capture. Enfin, les exhaustivités étaient variables selon les RCP réparties sur le territoire.

Cette étude a permis de donner quelques pistes pour améliorer l'exhaustivité des données de l'observatoire Revela13, notamment l'ajout en routine des données du PMSI et la nécessité d'un meilleur repérage des cas les plus âgés.

MOTS CLÉS : CANCERS, SURVEILLANCE, EXHAUSTIVITÉ, CAPTURE-RECAPTURE, RÉUNION DE CONCERTATION PLURIDISCIPLINAIRE.

Citation suggérée : Laurence 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.* Saint-Maurice : Santé publique France, 2018. 72 p. Disponible à partir de l'URL : www.santepubliquefrance.fr

ISBN-NET : 979-10-289-0471-5 - RÉALISÉ PAR LA DIRECTION DE LA COMMUNICATION, SANTÉ PUBLIQUE FRANCE - DÉPÔT LÉGAL : SE'PTEMBRE 2018

Abstract

Observatory of acute leukemia, kidney and bladder cancer (Revela13) among adults in Bouches-du-Rhône department

Completeness study, 2013-2014

The Revela13 observatory was implemented in the Bouches-du-Rhône department, on April 2012. The aim was to carry out spatial analyzes to identify clusters of these cancers in the study area. Data collection was based on cases reporting by urologists and hematologists and multidisciplinary team meeting (MDT) sheets from the cancer coordination centers. After 3 years, a capture-recapture study, over the 2013-2014 years, was conducted to assess the completeness of the Revela13 observatory by using hospitalizations data and histological data as complementary sources. A two-source analysis was performed for acute leukaemias and a three-source analysis was performed for urinary cancers.

The complementary sources allowed us to recover 15% of new incidental cases for urinary cancers and 20% for acute leukaemias, over the 2013-2014 years. The estimated number of cases was 232 [95% CI: 225-239] for acute leukaemias, 1247 [1239-1255] for bladder cancers and 600 [589-611] for kidney cancer. The overall completeness of the observatory was respectively 76% [74%-79%], 84% [83%-85%] and 84% [83%-86%]. The oldest people were significantly less likely to be identified by the Revela13 observatory than the younger ones. A catch heterogeneity related to the tumor stage and the residence's area was observed for bladder cancers only. Finally, the completeness was variable according to the cancer coordination centers.

The completeness of the Revela13 observatory would be improved by routinely inclusion of hospitalization data and by a better recording for the oldest aged cases.

KEY WORDS: CANCER, SURVEILLANCE, CAPTURE-RECAPTURE, COMPLETENESS

Rédaction du rapport

Laurence Pascal

Cellule d'intervention en région Provence-Alpes-Côte d'Azur (PACA) et en Corse (Cire PACA-Corse). Santé publique France

Karine Mantey

Observatoire régional d'épidémiologie en région Provence-Alpes-Côte d'Azur (Orep), Faculté de médecine, Marseille

Relecture

Jean-Claude Desenclos

Directeur scientifique de Santé publique France, Saint-Maurice, France

Remerciements

Pour leur participation au recueil des données et sans qui cette étude n'aurait pu se faire :

- Les responsables des départements d'information médicale des établissements publics et privés sollicités ;
- Les responsables des laboratoires d'anatomopathologie publics et privés sollicités ;
- Les réseaux régionaux OncoPaca-Corse et Onco-LR ;
- Les centres de coordination en cancérologie de l'AP-HM, de l'IPC, de Nîmes.

Pour leur implication dans l'étude, leur soutien inconditionnel et leur intérêt constant :

- Le professeur Karsenty, coordinateur du réseau des urologues ;
- Les professeurs Vey et Eisinger, coordinateurs du réseau des hématologues ;
- Les médecins responsables des différentes Réunions de concertations pluridisciplinaires participants à l'observatoire Revela13 ;
- Les urologues et les hématologues participants aux réseaux ainsi que les personnels des services hospitaliers et des cabinets médicaux pour leur accueil et leur aide.

Abréviations

ACP	Anatomo-cytopathologie
Adicap	Association pour le développement de l'informatique en cytologie et en anatomie pathologique
AIC	<i>Akaike Criteria Information</i>
AP-HM	Assistance publique des hôpitaux de Marseille
ARS	Agence régionale de santé
ASIP	Agence des systèmes d'information partagés de santé
BIC	<i>Bayesian Information Criteria</i>
CCEP	Comité consultatif d'évaluation des projets
CCTIRS	Comité consultatif sur le traitement de l'information en matière de recherche
CH	Centre hospitalier
CHRU	Centre hospitalier régional universitaire
CIM 10	Classification internationale des maladies de l'OMS 10 ^e révision
CIM O3	Classification internationale des maladies pour l'oncologie (3 ^e édition)
CIS	Carcinome <i>in-situ</i>
Cire	Cellule d'intervention de Santé publique France en région
CME	Commission médicale d'établissement
Cnil	Commission nationale informatique et libertés
CR	Compte rendu
DIM	Département de l'information médicale
DP	Diagnostic principal
DA	Diagnostic associé
DSIO	Département des systèmes d'information et de l'organisation
EOPMS	Espaces opérationnels de proximité médicaux-sociaux
HAS	Haute Autorité de santé
HP	Hôpital privé
IC	Intervalle de confiance
IPC	Institut Paoli-Calmette
InVS	Institut de veille sanitaire (devenu Santé publique France en mai 2016)
LAL	Leucémies aiguës lymphoblastiques
LAM	Leucémies aiguës myéloblastiques
Onco-LR	Centre régional de cancérologie en Languedoc-Roussillon
OncoPaca-Corse	Centre régional de cancérologie en Provence - Alpes - Côte d'Azur et Corse
Orep	Observatoire régional d'épidémiologie pour la région PACA
PMSI	Programme de médicalisation des systèmes d'information
PACA	Provence - Alpes - Côte d'Azur
RCP	Réunion de concertation pluridisciplinaire
RUM	Résumé d'unité médicale
Sp France	Santé publique France
TNM	Tumeur, Nodes, Metastasis
TVNIM	Tumeur vésicale n'infiltrant pas le muscle
TVIM	Tumeur vésicale infiltrant le muscle
UICC	Union internationale contre le cancer
3C	Centre de coordination en cancérologie

Sommaire

Abréviations	5
1. CONTEXTE ET OBJECTIFS	8
1.1 Contexte.....	8
1.2 L'observatoire Revela13 : une surveillance territorialisée de cancers ciblés	9
1.2.1 Lancement du projet (2011-2012).....	9
1.2.2 La faisabilité du recueil des données (2012-2014).....	9
1.2.3 Premières estimations de répartition territoriale	10
1.3 Objectifs	11
2. MATÉRIEL ET MÉTHODES	12
2.1 Modalités de recueil des données de l'Observatoire Revela13.....	12
2.1.1 Réseau des spécialistes partenaires	12
2.1.2 Structures partenaires pour les extractions de fiches RCP	13
2.2 Modalités de recueil des nouvelles sources de données	14
2.2.1 Données du PMSI	15
2.2.2 Données des laboratoires d'anatomo-cytopathologie	15
2.2.3 Données des laboratoires d'hématologie.....	16
2.2.4 Transmission des données.....	16
2.3 Application de la méthode capture-recapture aux données de l'Observatoire Revela13.....	17
2.3.1 Méthode capture-recapture	17
2.3.2 Constitution de la base de données.....	18
2.3.3 Connaissances <i>a priori</i> sur les facteurs d'hétérogénéité et de dépendance.....	19
2.3.4 Analyses des données	20
2.3.5 Outil d'analyse	21
3. RÉSULTATS.....	22
3.1 Cas identifiés par chacune des sources	22
3.1.1 Revela13	22
3.1.2 PMSI	24
3.1.3 Laboratoires ACP	26
3.2 Exhaustivité pour les leucémies aiguës	28
3.2.1 Description des données par sources.....	28
3.2.2 Analyse globale	29
3.2.3 Analyse stratifiée	29
3.2.4 Analyse par RCP d'hématologie	30
3.3 Exhaustivité pour les tumeurs urothéliales.....	31
3.3.1 Description des données par source.....	31
3.3.2 Dépendance entre sources - analyse globale	32
3.3.3 Analyse stratifiée	33
3.3.4 Analyse par RCP d'urologie.....	36
3.4 Exhaustivité des cancers du rein	36
3.4.1 Description des données par source.....	36
3.4.2 Dépendance entre sources - analyse globale	38
3.4.3 Analyse stratifiée	39
3.4.4 Analyse par RCP d'urologie.....	40

4. Discussion	41
4.1 Principaux résultats	41
4.2 Recueil des données	41
4.2.1 Extractions des fiches de RCP pour la complétude de l'Observatoire.....	41
4.2.2 Recueil des données des sources complémentaires	42
4.3 Méthode capture-recapture	43
4.3.1 Vérification des hypothèses	43
4.3.2 Indépendance des sources.....	44
4.3.3 Hétérogénéité de capture	44
4.4 Estimations de l'exhaustivité de l'Observatoire Revela13	45
4.4.1 Exhaustivité des cancers des voies urinaires.....	45
4.4.2 Exhaustivité des leucémies aiguës.....	46
5. Conclusions et perspectives	48
5.1 Amélioration de l'exhaustivité de l'observatoire Revela13.....	48
5.1.1 Optimiser les requêtes d'extraction des RCP.....	48
5.1.2 Compléter le recueil des données de RCP	49
5.1.3 Mieux cibler les cancers des personnes âgées.....	50
5.2 Compléter l'étude d'exhaustivité sur les années 2015-2016	51
Références bibliographiques	52
ANNEXES.....	54
Annexe 1 : Fiches de recueil	55
Annexe 2 : Libellés des codes CIM10.....	58
Annexe 3 : Liste des établissements hospitaliers et leur commune d'implantation.....	59
Annexe 4 : Liste des laboratoires d'anatomopathologies et laboratoires d'hématologies sollicités pour l'étude d'exhaustivité.....	60
Annexe 5 : Liste des codes Adicap.....	61
Annexe 6 : Schéma de croisement des données prévu dans le protocole	62
Annexe 7 - Rappel sur la méthode capture-recapture	63
Annexe 8 - R-capture	69
Annexe 9 : Analyse de sensibilité avec l'estimateur de Chao	71
Annexe 10 : Estimations du nombre de cas de tumeurs urothéliales et du taux d'exhaustivité des sources stratifiées sur chez les hommes et les femmes en fonction de leur classe d'âge et de leur zone de résidence, Bouches-du-Rhône, 2013-2014	72

1. CONTEXTE ET OBJECTIFS

1.1 Contexte

La surveillance des cancers se fait au travers d'un réseau de registres départementaux qui ont pour vocation un enregistrement exhaustif des cancers selon des modalités standardisées à l'échelle nationale et internationale. Ces registres permettent de produire des données en termes de caractéristiques des cancers, d'incidence, de prévalence et de survie et permettent la mise en œuvre de programmes de recherche. Un registre de cancers est un dispositif lourd et coûteux qui n'a pas vocation à être développé sur tout le territoire national.

Les données de ces registres sont utilisées par Santé publique France pour développer et valider les méthodes d'analyses des données des bases médico-administratives permettant d'estimer l'incidence des cancers au niveau national et infranational.

Ce système de surveillance ne permet pas de répondre aux problématiques locales, notamment aux préoccupations en santé-environnementale. La région Provence-Alpes-Côte d'Azur (PACA), et plus particulièrement le département des Bouches-du-Rhône sont largement concernés par ces problématiques et l'Agence régionale de santé (ARS) de PACA est régulièrement interpellée par des riverains, des médecins et des élus locaux sur l'impact sanitaire, notamment en termes de cancers, des niveaux de pollution autour de certains « points noirs environnementaux ».

En 2010, l'ARS PACA saisissait la Cellule d'intervention de Santé publique France en région (Cire) PACA-Corse sur la possibilité de mettre en place une surveillance des cancers au travers d'un dispositif adapté aux spécificités du territoire. L'enjeu était d'apporter des informations sur l'état de santé en termes de cancers à la population résidant dans ces zones particulièrement polluées. La Cire proposait alors de conduire une étude de faisabilité, à l'échelle du département des Bouches-du-Rhône, visant à disposer en continu de données sur des cas incidents de cancers, permettant une analyse territorialisée des données à un niveau géographique fin. Le dispositif retenu se devait d'être adapté au territoire, réactif et flexible en tenant compte des ressources disponibles et des moyens mis à disposition. Pour répondre à ces contraintes, le choix a été de mettre le praticien à l'origine des diagnostics au centre du dispositif en adossant l'observatoire Revela13 à des réseaux de spécialistes des pathologies ciblées et de focaliser le recueil sur une source de données permettant d'obtenir des informations sur la localisation des cas, sur la pathologie, sur l'histologie et le stade des cancers. Parmi les données disponibles au niveau local, les fiches dites « RCP », remplies pour chaque patient au moment de la réunion de concertation pluridisciplinaire (RCP) correspondaient le mieux à cet objectif. Les données issues du programme de médicalisation de l'information PMSI ne comprenaient en effet aucune donnée histologique et les données issues des laboratoires n'étaient pas facilement accessibles.

Les cancers ciblés pour cette expérimentation ont également été identifiés à partir des spécificités de ce département, en particuliers la présence d'une forte activité industrielle, et des connaissances sanitaires issus du ressenti de médecins locaux sur une augmentation croissante des cancers de la vessie ainsi que d'études antérieures menées par la Cire PACA-Corse.

En avril 2011, l'observatoire des cancers du rein, de la vessie et des leucémies aiguës chez l'adulte dans le département des Bouches-du-Rhône (Revela13) a été mis en œuvre. L'objectif principal de cet observatoire était la surveillance territorialisée de ces cancers incidents et notamment la détection d'éventuels agrégats spatio-temporels et/ou de zones de concentration de cancers. Cet observatoire permettait également de mettre à disposition des

autorités sanitaires des données locales validées, recueillies de façon homogène, dans un temps plus réduit et dans un cadre moins contraint que celui d'un registre de cancer classique.

L'étude de faisabilité, initialement programmée sur deux ans (avril 2011 - avril 2013) et prolongée d'une année supplémentaire (avril 2014), visait à tester les méthodes pour recenser de nouveaux cas de cancers sur le département des Bouches-du-Rhône, les géolocaliser et rechercher des regroupements spatiotemporels de cas, sans point source défini *a priori*.

1.2 L'observatoire Revela13 : une surveillance territorialisée de cancers ciblés

1.2.1 Lancement du projet (2011-2012)

La première année était consacrée à la rencontre des partenaires présents sur le territoire et au montage de l'observatoire. Une convention cadre tripartite entre l'ARS PACA, Santé publique France (ex InVS) et l'Observatoire régional d'épidémiologie de la région PACA (Orep) et deux conventions de partenariat avec l'Assistance publique des hôpitaux de Marseille (AP-HM) et l'Institut Paoli-Calmette (IPC), centre régional de référence du traitement du cancer étaient ainsi signées. Les autorisations légales étaient obtenues le 29 mars 2012 (autorisation Cnil DR-2012-138 du 21 mars 2012).

Le recensement des cas reposait sur des réseaux de médecins spécialistes, à l'origine du diagnostic des cancers ciblés. L'ensemble des médecins urologues et des hématologues du département ainsi que des principaux centres de diagnostic des départements limitrophes, susceptibles de recevoir des patients des Bouches-du-Rhône, avaient été contactés. L'équipe avait également visité l'ensemble des RCP spécifiques et polyvalentes identifiées pour présenter le projet et réunir les médecins spécialistes au sein d'un réseau dédié à l'observatoire. La création et la coordination de ces réseaux étaient assurées conjointement par l'équipe projet et par des médecins experts désignés coordonnateurs de réseaux de spécialistes. Chaque médecin participant au réseau signait une charte d'adhésion. Deux urologues avaient refusé d'y participer, l'un exerçait à Marseille et l'autre à Martigues.

Une base de données dédiée était créée à partir de l'application Voozаноо, gérée par Epiconcept®, et était localisée à Santé publique France (SpFrance). Elle était accessible par internet et chaque adhérent qui le souhaitait pouvait saisir ses cas en ligne avec un mot de passe personnel transmis par l'équipe projet. En pratique, cette fonctionnalité a peu été utilisée et les fiches papiers étaient saisies par l'équipe projet.

1.2.2 La faisabilité du recueil des données (2012-2014)

Le recueil des données démarrait le 2 avril 2012. Il était basé sur une surveillance prospective, à partir du signalement par le médecin à l'origine du diagnostic, de tout nouveau cas de cancer des voies urinaires ou d'une leucémie aiguë chez un patient adulte résidant dans le département des Bouches-du-Rhône. La complétude de ces signalements directs était assurée par des extractions trimestrielles des fiches dites « RCP ». En effet, depuis le 1^{er} plan Cancer 2003-2007, chaque nouveau patient atteint de cancer doit bénéficier d'une RCP autour de son dossier à l'issue de laquelle une fiche RCP est remplie puis saisie sur une base de données dédiée. Les fiches RCP sont ensuite adressées au centre de coordination en cancérologie (3C) de l'établissement de référence puis centralisées au niveau du réseau régional de cancérologie. Des extractions trimestrielles de la base de données RCP de ce réseau pour la région PACA, le réseau régional OncoPaca-Corse, permettaient à l'équipe projet d'identifier des patients qui n'auraient pas déjà été signalés à l'observatoire. D'autres extractions, issues des 3C des établissements, non connectés au réseau régional (AP-HM et

IPC), ou des 3C des départements limitrophes (Gard, Vaucluse), étaient également régulièrement transmises. Les données complémentaires du patient étaient ensuite collectées au travers d'une fiche de signalement adaptée à chaque cancer auprès des médecins en charge du patient (Annexe 1).

Chaque cas validé était ensuite enregistré sur la base de données Revela13. Lors de la saisie des coordonnées du patient, les cas étaient géolocalisés automatiquement à l'adresse de résidence au moment du diagnostic.

La visite des RCP concernées par le projet et la rencontre des médecins spécialistes avaient permis de mettre en place des modalités de recueil de données adaptées à chaque RCP et à chaque praticien, pour les médecins qui ne participaient pas encore régulièrement à des RCP. Ces modalités étaient modifiées en fonction des départs et arrivées de ces spécialistes et/ou des changements dans l'adhésion à une RCP. De même, l'organisation des réseaux à partir des RCP permettait à l'équipe projet d'informer et de sensibiliser régulièrement ces spécialistes.

La faisabilité d'enregistrement des cas était démontrée dès 2014, avec un nombre de cas observé globalement supérieur au nombre de cas attendu, estimé à partir des incidences nationales pour chaque cancer (750 cas par an pour les 3 cancers). Le nombre de cas observé était notamment plus élevé pour les cancers urothéliaux, identique pour les leucémies aiguës et plus faible pour les cancers du rein. Ainsi, entre le 2 avril 2012 et le 31 décembre 2014, 1 392 cas de tumeurs urothéliales, 689 cancers du rein et 252 cas de leucémies aiguës étaient inclus dans l'observatoire.

1.2.3 Premières estimations de répartition territoriale

Les premières cartographies réalisées en 2013 à l'échelle des espaces opérationnels de proximité médico-sociaux (EOPMS) définis par l'ARS, reflétant bien la répartition des groupes de RCP avec lesquels l'équipe travaille, mettaient en évidence une hétérogénéité dans la distribution spatiale des cas au sein du département des Bouches-du-Rhône. Les taux d'incidence calculés étaient ainsi plus bas à l'ouest du département pour les trois cancers étudiés. Ces cartes avaient suscité beaucoup de discussions avec les partenaires de l'observatoire sur les éventuelles fuites de patients vers les départements du Gard et de l'Hérault mais également sur une sur incidence liée à une forte activité et une offre de soins plus importante dans la partie est du département.

Le Comité technique du projet, consulté en octobre 2013, jugeait nécessaire de vérifier l'hypothèse d'une fuite des patients du département des Bouches-du-Rhône vers les départements limitrophes du Gard, de l'Hérault et du Vaucluse. Cette démarche devait également être complétée par une enquête d'exhaustivité, comme prévu dans le protocole d'étude initial et rappelé par le Comité collégial d'évaluation des projets (CCEP) de SpFrance, dès le mois de septembre 2012. La connaissance de l'exhaustivité du système avait également été identifiée comme indispensable au développement d'analyses spatiales, que ce soit pour la réalisation de cartes de risques ou pour les méthodes de *clustering*, par le comité scientifique interne à Santé publique France en juin 2013.

Un travail sur une éventuelle fuite de patients de l'ouest du département vers les départements limitrophes, au travers des extractions des 3C du Gard et du Vaucluse, et, de la mise en place d'une nouvelle collaboration avec le réseau régional de cancérologie du Languedoc-Roussillon (réseau Onco-LR), était rapidement mis en œuvre. Si quelques cas avaient pu être récupérés, ils n'expliquaient pas la faible incidence observée dans la partie ouest du département. Parallèlement, une demande d'avenant à l'autorisation Cnil du projet était déposée dès février 2014 pour mettre en œuvre l'étude d'exhaustivité à partir de nouvelles sources de données.

L'accord Cnil était obtenu en octobre 2014 (autorisation DR-2014-442 du 8 octobre 2014) et l'enquête d'exhaustivité démarrait en 2015.

1.3 Objectifs

L'étude d'exhaustivité avait pour objectif d'estimer l'exhaustivité globale de l'observatoire pour les trois premières années d'enregistrement des cas (2012 à 2014), et plus spécifiquement, d'étudier l'exhaustivité de chacune des RCP déclarant des cas de cancers urologiques ou hématologiques. L'exhaustivité de l'observatoire était mesurée à partir d'une étude capture-recapture à trois sources : la base de données Revela13, les données du Programme de médicalisation des systèmes d'informations (PMSI), produites par les départements de l'information médicale (DIM) des établissements qui prennent en charge des cas des trois cancers ciblés, et les données issues des laboratoires d'anatomo-cytopathologies (ACP) et d'hématologie à l'origine des diagnostics biologiques de cancers.

2. MATÉRIEL ET MÉTHODES

L'objectif de cette étude rétrospective était d'estimer l'exhaustivité globale de l'observatoire Revela13 pour les années 2013-2014 à partir d'une méthode capture-recapture à deux ou trois sources. Le recueil de l'année 2012 ayant démarré au 2^{ème} trimestre, elle n'a pas été incluse dans les analyses même si les données recueillies sont présentées dans la partie des résultats concernant l'apport des sources complémentaires.

2.1 Modalités de recueil des données de l'observatoire Revela13

Le recueil de données de l'observatoire Revela13 était basé, au début, sur le signalement direct des médecins spécialistes à l'origine des diagnostics de cancers, complété par des extractions trimestrielles des fiches RCP. Très rapidement, les signalements via les fiches RCP sont devenus la principale source de données de l'observatoire Revela13. En région PACA, les fiches RCP sont centralisées par le réseau régional OncoPaca-Corse. Cependant, certains 3C ne transmettent pas encore leurs fiches au réseau régional, il s'agit des 3C de l'AP-HM et de l'IPC. Une requête d'extraction a été créée pour chaque structure partenaire.

2.1.1 Réseau des spécialistes partenaires

Lors de la mise en place de l'observatoire, 48 chirurgiens-urologues et 18 onco-hématologues avaient été identifiés sur le département des Bouches-du-Rhône. Parmi eux, 44 urologues et 14 hématologues avaient accepté d'adhérer au réseau de l'observatoire en signant une charte d'adhésion. Par cette adhésion, les médecins spécialistes s'engageaient à permettre à l'équipe projet de compléter les signalements de patients ayant un premier diagnostic de l'un des cancers retenus, et à informer ces patients de leur inclusion dans l'observatoire au moment qui leur semblait opportun.

Les chirurgiens-urologues participaient pour la plupart d'entre eux à l'une des neuf RCP d'urologie du département. Les autres étaient associés à une RCP polyvalente (dont l'une avec une forte orientation en urologie). Les 4 urologues qui n'avaient pas adhéré au réseau n'étaient pas rattachés à une RCP particulière. Par ailleurs, les chirurgiens urologues de trois RCP d'urologie dans les départements limitrophes adhéraient également aux réseaux de l'observatoire : 2 RCP dans le Gard et 1 RCP dans le Vaucluse.

Au niveau du département, les onco-hématologues contactés participaient aux 4 RCP spécifiques en hématologie, l'une de ces RCP était rattachée à l'expertise de la RCP d'hématologie de l'IPC. Les spécialistes de la RCP d'hématologie du CH d'Avignon avaient également accepté de participer au réseau.

L'ensemble de ces médecins spécialistes avaient la possibilité de signaler directement un cas à l'observatoire par fax, mail ou courrier ou encore en saisissant directement en ligne le signalement sur la base de données Voozаноо. En effet, les spécialistes pouvaient bénéficier d'un accès à la base de données en ligne après obtention d'un identifiant et d'un mot de passe personnel. Cette dernière modalité ayant été très peu utilisée, elle n'avait pas été conservée à partir de 2015.

2.1.2 Structures partenaires pour les extractions de fiches RCP

Réseau régional OncoPaca-Corse

Lors de la mise en place de l'observatoire Revela13, la requête d'extraction des fiches RCP des patients susceptibles de répondre aux définitions de cas retenues par l'observatoire avait été élaborée avec le réseau régional OncoPaca-Corse. Cette requête était basée sur tout nouveau diagnostic de cancer avec un code de la classification internationale des maladies 10^e révision (CIM10) [1] prédéfini (Annexe 2), chez un adulte résidant dans les Bouches-du-Rhône ou sans domicile précisé, et dont le diagnostic avait été discuté dans une RCP où les cas de cancers urinaires ou de leucémies aiguës pouvaient être présentés. Les variables retenues dans la requête étaient les suivantes : nom du patient et nom marital, prénom, date de naissance, sexe, date de début de la maladie, date de prélèvement, stade pTNM de la tumeur, lieu de résidence du patient, code CIM10, date et intitulé de la RCP, établissement de prise en charge, noms et coordonnées des médecins généralistes et spécialistes référents.

Entre 2012 et 2014, la transmission des RCP au réseau régional OncoPaca-Corse a beaucoup évolué avec l'intégration progressive des RCP de l'hôpital Saint-Joseph début 2013 puis, un peu plus tard, du 3C du Vaucluse. Durant cette période, les transmissions automatiques des 3C vers le réseau régional ont également connu de nombreuses interruptions, notamment pour la RCP d'urologie de la Polyclinique du Parc Rambot ou encore celle du centre hospitalier du Pays d'Aix. Ces interruptions pouvaient être dues à des problèmes informatiques ou à l'absence de personnel remplissant les fiches informatisées de RCP. Parallèlement, au fur et à mesure du déploiement de l'observatoire Revela13, une meilleure connaissance des dossiers discutés en RCP permettait d'améliorer, en lien avec l'équipe du réseau régional, la détection des cas incidents de cancers en modifiant les requêtes d'extraction. Ainsi, avant même le début de l'étude d'exhaustivité, deux modifications majeures de la requête d'extraction du réseau régional avaient été mises en œuvre.

En janvier 2014, une première modification était réalisée avec l'ajout de nouveaux codes CIM10 : les codes de la série D41* pour identifier les cancers urologiques en cours d'investigation, hors prostate, et les codes C835 et C837 pour mieux identifier les lymphomes lymphoblastiques et les lymphomes de Burkitt. En janvier 2015, au vu de cas de cancers identifiés lors des passages dans les établissements de santé et non détectés dans les extractions de RCP, la requête d'extraction était modifiée de manière à prendre en compte l'ensemble des dossiers patients discutés en RCP et non plus uniquement les dossiers pour lesquels un premier passage en RCP avait été noté. Ainsi un patient dont le dossier était passé une première fois en RCP pour une suspicion de cancer non confirmé histologiquement pouvait être récupéré dans les extractions lors d'un second passage en RCP. La nouvelle requête proposée permettait ainsi d'élargir la détection des patients correspondant potentiellement à la définition de cas de l'observatoire.

Centre de coordination en cancérologie de l'AP-HM

Concernant le 3C de l'AP-HM, une requête d'extraction était élaborée à partir des définitions de cas retenues par l'observatoire Revela13 pour chacune des trois RCP spécifiques identifiées : les RCP d'urologie et d'hématologie de l'hôpital de la Conception et la RCP d'urologie de l'hôpital Nord. Ces extractions concernaient tout nouveau cas de cancer dont le dossier avait été discuté dans l'une de ces trois RCP. Cette requête était modifiée en 2015 pour ajouter les codes CIM10 (D41*, C835, C837). Les variables retenues dans la requête étaient les suivantes : date et numéro de la RCP, nom marital et nom de naissance du patient, prénom, sexe, date de naissance, code postal et commune de résidence, diagnostic en code CIM10, date de prélèvement, type histologique, stade pTNM et grade de la tumeur.

Centre de coordination en cancérologie de l'IPC

La requête a été élaborée par l'équipe du département des systèmes d'information et de l'organisation (DSIO), à partir des définitions de cas retenues par l'observatoire. Les variables retenues dans la requête étaient les suivantes : nom marital et nom de naissance du patient, prénom, sexe, date de naissance, adresse, code postal et ville de résidence, diagnostic en code CIM10, intitulé du diagnostic principal, date de diagnostic, date et intitulé de la RCP, nom et prénom du médecin traitant.

En avril 2013, l'IPC ayant changé de système d'exploitation, la stabilité du requêtage a été impactée, et jusqu'en 2014, les requêtes n'avaient pas de format standard. À partir de janvier 2015, un nouveau travail sur les requêtes était effectué, d'une part pour ajouter les codes CIM10 (D41*, C835, C837) identifiés comme nécessaires à l'amélioration de l'identification des cas et, d'autre part, pour prendre en compte les patients dont les cancers n'étaient pas confirmés histologiquement. Deux requêtes complémentaires étaient ainsi élaborées : une extraction par code CIM10 et une extraction par RCP spécialisée (RCP urologie, RCP hématologie et RCP lymphome).

Centre de coordination en cancérologie de Nîmes

Une extraction semestrielle était adressée à l'observatoire par les 3C Public et Privés de Nîmes. Ces extractions portaient uniquement sur les patients résidant dans le département des Bouches-du-Rhône. Les variables extraites étaient issues du protocole d'extraction mis en place avec le réseau régional OncoPaca-Corse.

Réseau régional Onco-LR

Une extraction annuelle des données du réseau Onco-LR permettait de compléter les données du 3C de Nîmes. En effet, quelques rares patients étaient diagnostiqués et suivis dans les structures publiques et privées de Montpellier.

2.2 Modalités de recueil des nouvelles sources de données

Les sources complémentaires de données étaient :

- les données du PMSI qui identifient les passages des patients dans une structure de soin pour la prise en charge des trois cancers ciblés ;
- les comptes rendus (CR) d'anatomopathologie ou de cytologie des laboratoires d'ACP qui réalisent des cytologies urinaires, correspondant à une suspicion de tumeurs malignes, et des histologies de cancers invasifs urologiques, hors prostate ;
- les résultats des myélogrammes des laboratoires d'hématologie qui attestent la survenue d'une leucémie aiguë chez un patient.

Les données du PMSI provenaient des DIM des établissements publics et privés ayant déjà reçu au moins un cas de l'un des cancers ciblés pendant la période d'étude, que ce soit dans le département des Bouches-du-Rhône ou dans l'un des départements limitrophes. La liste des établissements sollicités est présentée en Annexe 3.

Les laboratoires d'ACP et d'hématologie retenus pour l'étude étaient identifiés à partir des laboratoires ayant déjà réalisés au moins un diagnostic de cancer des voies urinaires ou de leucémie aiguë chez un patient inclus dans la base de données de l'observatoire Revela13 (Annexe 4).

2.2.1 Données du PMSI

Vingt-neuf établissements de santé ont été sollicités pour réaliser une extraction des données de leur base PMSI : 8 centres hospitaliers publics dont deux dans des départements limitrophes (CH d'Avignon et centre régional hospitalo-universitaire (CHRU) de Nîmes), 6 hôpitaux privés, 14 cliniques privées, dont la clinique Rhône-Durance à Avignon, la polyclinique Kenval à Nîmes, ainsi que l'Institut Paoli Calmette, centre régional de lutte contre le cancer.

L'ensemble des directeurs, des présidents de la commission médicale d'établissement (CME) et des responsables des DIM des 29 établissements partenaires ont été informés par courrier puis par messagerie de la mise en place d'une analyse des données du PMSI dans le cadre de cette étude. Une rencontre préalable avec chaque responsable de DIM a été proposée pour discuter des modalités d'extractions et de transmission des données.

Les variables à extraire étaient les suivantes : nom, nom de jeune fille, prénom, date de naissance, sexe, code postal du lieu de résidence, numéro d'hospitalisation, numéro d'unité fonctionnelle, nom du chef de service, diagnostic principal, dates d'entrée et de sortie.

Les séjours extraits des fichiers de Résumés d'unité médicale (RUM) ont été identifiés à partir des diagnostics principaux (DP), correspondant à liste de codes CIM10 retenus (Annexe 2) et d'un code postal de résidence dans les Bouches-du-Rhône. Les données ont été extraites sur la période 2010-2014. Les deux premières années ont permis d'identifier les cas non incidents sur les années 2012 à 2014, en excluant tous les patients déjà hospitalisés pour l'un des cancers étudiés sauf pour les cancers de la vessie. Les modalités d'extractions et de formats ont été adaptées à chaque établissement.

La durée entre l'envoi de la demande et l'obtention des données était variable, de quelques semaines à presque 2 ans, avec pour certaines structures, nécessité de plusieurs relances. Certains médecins DIM, souvent sur plusieurs établissements, avaient peu de temps pour participer à une enquête épidémiologique. Parfois les extractions ont été réalisées par des informaticiens, des gestionnaires de données ou des prestataires externes.

L'ensemble des structures sollicitées avait transmis leurs données sous format informatique, ce qui a permis de croiser les fichiers avec la base de données Revela13. Seuls les doublons de cas déjà inclus ou nouvellement inclus ont été identifiés, les doublons de cas exclus n'ont pas été comptabilisés.

2.2.2 Données des laboratoires d'anatomo-cytopathologie

Concernant les données anatomo-cytopathologiques, 17 laboratoires ont été contactés : 12 sur le département des Bouches-du-Rhône et 5 sur les départements limitrophes (3 dans le Gard et 2 dans le Vaucluse). Parmi ces laboratoires, 6 appartenaient à des structures hospitalières et 11 étaient des laboratoires privés (Annexe 4).

Dans un premier temps, une rapide enquête téléphonique a permis de mieux connaître les pratiques de codage des CR d'anatomo-pathologie et le logiciel informatique utilisé dans les principaux laboratoires. Un courrier a été ensuite adressé à l'ensemble des laboratoires en juillet 2015 pour présenter l'étude. Celui-ci était accompagné d'un listing de codes CIM10 (Annexe 2) et de codes Adicap [2] (Annexe 5), identifiant les cytologies urinaires correspondant à une suspicion de tumeurs malignes, les histologies de tumeurs urothéliales invasives et des tumeurs malignes du rein retenus dans l'étude. Un travail spécifique sur les codes de la classification Snowmed a été réalisé pour un des laboratoires mais n'a finalement pas été utilisé, l'extraction informatisée s'avérant trop compliquée.

Les variables à extraire étaient les suivantes : noms, prénom, date de naissance (ou année de naissance), sexe, code postal du lieu de résidence, code Adicap ou code CIM10, diagnostic principal, médecin prescripteur et établissement dans lequel l'acte avait été réalisé. Les modalités d'extractions ont été adaptées à chaque laboratoire, l'extraction d'un fichier informatisé étant privilégiée.

L'ensemble des laboratoires du département des Bouches-du-Rhône (5 hospitaliers et 7 privés) ont participé à l'étude. Seul un laboratoire privé sur les deux du Vaucluse a accepté d'y participer ; 2 laboratoires privés et un laboratoire public dans le Gard ont refusé.

La durée entre l'envoi de la demande et l'obtention des données a été variable, de quelques semaines à dix-huit mois. Pour certaines structures plusieurs relances ont été nécessaires et des rendez-vous sur site pour voir le logiciel et trouver la solution la mieux adaptée ont souvent été mises en œuvre.

Sur les 13 laboratoires participants, 11 avaient réalisé une extraction de leur base de données à partir des codes Adicap. Cette extraction a été transmise sous format informatique pour 6 d'entre eux et sous format papier pour les 5 autres. Pour deux laboratoires n'utilisant ni la codification Adicap ni la codification CIM10, les extractions ont concerné l'ensemble des CR anatomopathologiques des urologues avec lesquels ils avaient travaillé au moins une fois. Les CR papiers ont été ensuite triés manuellement pour retenir les CR correspondant aux tumeurs urothéliales et aux cancers du rein.

2.2.3 Données des laboratoires d'hématologie

Neuf laboratoires d'hématologie ont été identifiés dont deux hors du département des Bouches-du-Rhône (CH d'Avignon et CHRU de Nîmes) (Annexe 4). Tous les laboratoires avaient été contactés par courrier ou messagerie. Le laboratoire de l'AP-HM était en mesure de réaliser une extraction de myélogrammes correspondant à un diagnostic de leucémie aiguë puisqu'il transmettait ces données au 3C. Le laboratoire de l'IPC, des hôpitaux d'Aix-en-Provence et de Martigues n'ont pas été en mesure de fournir ces données et seule une extraction de tous les myélogrammes sur les trois années d'étude était possible. Le laboratoire du CH d'Avignon devait faire une requête avec le prestataire externe mais les données n'ont jamais été envoyées. Les autres laboratoires n'ont pas été enquêtés compte-tenu de l'impossibilité d'obtenir les données des principaux laboratoires. Cette source de données n'a donc pas été exploitée.

2.2.4 Transmission des données

La transmission sécurisée des fichiers de données nominatives a été assurée par leur dépôt sur Epifiles, une plateforme sécurisée de partage des données de santé gérée par Epiconcept® et agréée par l'Agence des systèmes d'information partagés des données de santé (ASIPsanté). Chaque partenaire disposait d'un identifiant unique et d'un mot de passe, qu'il définissait lui-même à la première connexion. Les fichiers de données étaient déposés dans un entrepôt spécifique à l'étude d'exhaustivité. Les fichiers récupérés par l'équipe projet étaient ensuite supprimés de la plateforme.

Pour l'AP-HM, un travail préalable avec le conseiller informatique et libertés (CIL) et le responsable de la sécurité informatique a été nécessaire pour obtenir l'autorisation de transmettre des données nominatives via Epifiles.

2.3 Application de la méthode capture-recapture aux données de l'observatoire Revela13

2.3.1 Méthode capture-recapture

La méthode capture-recapture permet, en croisant plusieurs sources d'information issues d'une même population, d'estimer le nombre de cas qui ne sont identifiés par aucune des sources, le nombre de cas total (N) de la maladie et ainsi l'exhaustivité de chaque source [3]. Initialement utilisée en zoologie et en démographie, cette technique est aujourd'hui régulièrement appliquée à l'épidémiologie pour estimer l'exhaustivité d'un système de surveillance [4,5], d'un registre de cancers [6] ou de morbidité [7].

Les principes, conditions d'application et limites de cette méthode ont été décrits dans la littérature [8-10] et sont présentés plus en détail en Annexe 7.

Principe de la méthode capture-recapture

Cette méthode permet, en croisant différentes sources et après avoir identifié les cas communs, d'estimer le nombre de cas présents dans aucune des sources. Il est au final possible d'estimer le nombre total de cas présents dans la population d'étude ainsi qu'un intervalle de confiance.

Conditions d'application

La fiabilité des résultats dépend du respect de différentes conditions d'applications :

- Identification de vrais cas ;
- Identification de tous les vrais cas communs ;
- Même période et même zone d'étude pour toutes les sources ;
- Homogénéité de capture ;
- Indépendance entre les sources ;
- Population d'étude close.

Le non-respect de ces conditions d'application peut conduire, selon les cas, à une surestimation ou une sous-estimation du nombre total de cas.

Prise en compte de la dépendance entre sources

Dans le cas d'un modèle à deux sources, l'estimation du nombre de cas est possible sous l'hypothèse d'indépendance des sources. Dans le cas d'un modèle à trois sources, la dépendance entre sources prises 2 à 2 peut être contrôlée en appliquant un modèle log-linéaire.

Prise en compte de l'hétérogénéité de capture

Il peut exister une hétérogénéité de capture liée à certaines caractéristiques des cas qui modifient leur probabilité de capture. Cette situation est fréquente lorsque l'on s'intéresse à des cas de cancers. Ce type d'hétérogénéité est identifié par des modèles stratifiés sur la ou les variables d'intérêts de manière à analyser des sous-groupes de cas avec une probabilité de capture plus homogène.

2.3.2 Constitution de la base de données

Le délai d'obtention des données ainsi que la diversité des fichiers informatiques récupérés ou l'absence de fichiers informatiques pour certaines sources (extraction de données transmise en format papier) n'a pas permis de réaliser un fichier global pour chacune des sources, comme cela était initialement prévu (Annexe 6).

Définition de cas commune aux 3 sources

La population cible était constituée de toute personne adulte (âgée de 18 ans et plus), résidant dans le département des Bouches-du-Rhône et pour laquelle un premier diagnostic de cancer du rein, de la vessie ou de leucémies aiguës avait été réalisé entre le 2 avril 2012 et le 31 décembre 2014. Les définitions de cas retenues sont communes aux trois sources.

Cas incident de leucémie aiguë

Un nouveau cas de leucémie aiguë correspond à un premier diagnostic de leucémie 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. Le diagnostic est basé sur différents examens biologiques incluant le myélogramme, l'immunophénotypage des blastes et les tests génétiques qui permettent de classer les cas en leucémies aiguës myéloïdes (LAM) ou lymphoïdes (LAL) [11].

Cas incident de cancer urothélial

Un nouveau cas de cancer urothélial correspond à un premier diagnostic de tumeur invasive de la vessie (CIM10 : C67), de l'uretère (C66), du bassinnet (C65) ou de l'urètre ou autres localisations (C68). Les tumeurs retenues dans l'étude sont les tumeurs classées pT1 à pT4 de la classification *Tumour, Nodes, Metastasis* (TNM) de l'Union internationale contre le cancer (UICC) [12].

Les résultats anatomopathologiques permettent de classer les tumeurs en pTa (tumeur vésicale sans franchissement de la membrane basale), pTis (carcinome in situ), pT1 tumeur ayant franchi la membrane basale mais limitée au chorion), pT2 (tumeur infiltrant le muscle), pT3 et pT4 (tumeur infiltrant les organes ou structures de voisinage) [13]. Mais en pratique, les urologues distinguent les tumeurs de vessie n'envahissant pas le muscle (TVNIM) regroupant les stades pTis, pTa et pT1 et les tumeurs de vessie infiltrant le muscle (TVIM) regroupant les stades pT2 à pT4.

L'incidence est calculée uniquement pour les cancers de la vessie.

Cas incident de cancer du rein

Un nouveau cas de cancer du rein correspond à un premier diagnostic de tumeur invasive du rein (hors voie excrétrice supérieure) correspondant au code CIM10 C64, les tumeurs retenues sont classées pT1 à pT4 dans la classification pTNM. La classification est basée sur la taille de la tumeur (pT1 : tumeur \leq 7 cm, pT2 : tumeur $>$ 7 cm limitée au rein), l'envahissement local (pT3 : tumeur avec thrombus veineux ou infiltrant le tissu adipeux, pT4 : tumeur envahissant le fascia de Gérota et/ou la surrénale), la présence ou non d'un envahissement ganglionnaire et la présence ou non de métastases [13].

Les cas de cancers retenus pour calculer l'incidence correspondent aux codes C64-C65-C66 et C68 de la CIM10 qui regroupent les adénocarcinomes du parenchyme rénal et les tumeurs malignes des voies excrétrices à l'exception de la vessie.

Identification des vrais cas

Pour chaque nouveau cas identifié dans un fichier issu d'un DIM ou d'un laboratoire, les critères d'inclusion ont été ensuite vérifiés auprès de l'établissement et du médecin prescripteur pour inclure de vrais cas.

Identification de la source de donnée

Les nouveaux cas validés ont été saisis dans la base de données Revela13. Une variable dédiée à l'enquête avait été ajoutée à la base (variable « exhaustivité »), elle était vide par défaut et codé 1 dès lors que le cas inclus était identifié par au moins une des deux sources complémentaires (PMSI ou ACP). L'origine du signalement a également été notifiée dans une variable dédiée « autre_source ».

Si le cas était identifié par plusieurs sources (inter-ACP et/ou inter-PMSI), il n'était saisi qu'une fois et toutes les sources qui avaient signalées ce cas étaient notifiées dans la variable dédiée.

Identification des doublons au sein de chaque source

Pour chaque fichier reçu, la première étape a consisté à éliminer les doublons au sein d'une même source. En effet, les patients atteints d'un cancer passaient souvent plusieurs fois dans le même établissement pour le suivi de leur maladie. De même, ils bénéficiaient souvent de plusieurs prélèvements analysés dans les laboratoires ACP.

Identification des cas communs aux différentes sources

Chaque nouveau fichier dédoublonné a été ensuite comparé au fichier initial de Revela13. Les cas communs ont été identifiés par croisement des fichiers des DIM et des laboratoires ACP après attribution du numéro unique d'identification utilisé dans la base Revela13 ou manuellement pour les fichiers papiers. Les cas non communs de chacune des sources complémentaires ont fait l'objet d'une vérification manuelle des cas de manière à prendre en compte les nombreuses erreurs de saisie (sexe et/ou date de naissance erronés, enregistrement des cas sous le nom de jeune fille pour les femmes, par exemple).

Constitution d'une base de données agrégée

Un fichier agrégé correspondant au format d'analyse du logiciel R-Capture a été construit en créant les variables Revela13, PMSI, ACP recodées en présence (1) ou absence des cas (0) dans les 3 sources respectives. Les sources ont été ensuite précisées dans deux nouvelles variables : NOM_DIM correspondant au libellé du ou des DIM sources et NOM_ACP correspondant au libellé du ou des ACP sources.

2.3.3 Connaissances *a priori* sur les facteurs d'hétérogénéité et de dépendance

Les signalements via les extractions de fiches RCP des différents partenaires étant devenus, dès 2013, la principale et quasi unique source de données de l'observatoire, la mesure de l'exhaustivité de la base de données Revela13 était le reflet de l'exhaustivité des RCP du département. De même, les limites et faiblesses des RCP identifiées dans la littérature risquaient d'être celles retrouvées pour l'observatoire.

Facteurs d'hétérogénéité

Les variables pour lesquelles une hétérogénéité de capture pouvait être suspectée ont été identifiées à partir de la littérature [4, 7, 14]. L'âge et le sexe des patients ainsi que le stade de la tumeur pour les cancers urinaires étaient retenus. Par ailleurs, avec la mise en œuvre

progressive de l'observatoire Revela13 et des stratégies de recueil de données adaptées à chaque RCP, l'année de diagnostic et le secteur de résidence des cas (les signalements directs par les médecins spécialistes étant plus fréquents sur Marseille) ont également été retenus comme facteurs potentiels d'hétérogénéité.

Pour les LAM, trois classes d'âge (moins de 60 ans, de 60 à 79 ans et plus de 80 ans) ont été construites à partir des recommandations de prise en charge de cas de leucémies aiguës différentes pour ces trois classes d'âge [15].

Concernant les deux cancers urinaires, ce sont les distributions des cas par âge qui ont permis de déterminer des classes d'âge. Pour les tumeurs urothéliales, les trois classes d'âge ont été : les moins de 65 ans, les 65-84 ans et les 85 ans et plus, et pour les cancers du rein : les moins de 55 ans, les 55-74 ans et les 75 ans et plus.

Le stade tumoral pour les cancers urinaires a été défini en deux classes « stade pT1 » et « stade >pT1 ».

L'année de diagnostic a été construite à partir de la date du diagnostic enregistrée pour chaque cas.

Le secteur de résidence a été défini à partir d'un regroupement des codes de résidence des patients en deux classes « résidant à Marseille » et « résidant Hors Marseille ».

Dépendance entre sources

Concernant la dépendance entre les sources, la codification d'un cas de cancer dans le PMSI étant très souvent liée au résultat de sa confirmation anatomopathologique, une dépendance entre les sources PMSI et ACP était donc vraisemblable.

De même, un cas n'est souvent présenté en RCP qu'après confirmation du diagnostic par le résultat de l'histologie, une dépendance positive devrait donc exister entre les sources Revela13 et ACP.

Une dépendance entre les sources Revela13 et PMSI pouvait également exister si la complétude des données de RCP était effectuée avec les données du PMSI.

2.3.4 Analyses des données

Représentativité des sources

La représentativité des sources selon les variables d'hétérogénéité a été évaluée par un test du Chi2.

Analyse globale

Une analyse globale de l'ensemble des données, réparties dans des tableaux de contingence récapitulant le nombre de cas en fonction de leur présence ou non dans chacune des sources, a été réalisée pour chacun des trois cancers étudiés.

L'analyse des leucémies aiguës a été réalisée par la méthode à deux sources puisque les données des laboratoires d'hématologie n'avaient pas pu être obtenues. L'analyse stratifiée sur les facteurs d'hétérogénéité a été réalisée avec un modèle log-linéaire.

Pour les tumeurs urothéliales et le cancer du rein, après estimation de la dépendance entre sources par la méthode de Wittes [16], les analyses ont été réalisées par modèle log-linéaire

avec l'estimateur de Poisson. Après ajustement, chaque modèle a conduit à une estimation du nombre total de cas (N estimé) et à son écart type, permettant de calculer les intervalles de confiance de N. Le modèle retenu pour estimer N était celui qui minimisait la déviance, le critère d'Akaike (AIC) et le critère d'information bayésien (BIC). Les intervalles de confiance ont été calculés par la méthode de Bishop [17].

L'exhaustivité de l'observatoire Revela13 a été estimée en rapportant le nombre de cas enregistrés dans la base Revela13 avant l'étude d'exhaustivité, au nombre total de cas estimés par le modèle sélectionné ainsi qu'aux bornes supérieure et inférieure de son IC à 95%. L'exhaustivité a été également estimée pour les sources PMSI et ACP ainsi que pour l'ensemble des 3 sources.

Une analyse de sensibilité utilisant l'estimateur de Chao a été réalisée. Cet estimateur permettait de contrôler en partie l'hétérogénéité et conduisait souvent à un intervalle de confiance moins large que l'estimateur de Poisson.

Analyse stratifiée

Pour étudier l'hétérogénéité de capture, des analyses log-linéaires avec estimateurs de Poisson ont été réalisées pour chacune des strates des variables d'hétérogénéité identifiées *a priori*. Le modèle appliqué était celui précédemment identifié dans l'analyse globale. L'exhaustivité ajustée de la source Revela13 a été estimée en rapportant le nombre de cas enregistrés dans la base Revela13 avant l'étude d'exhaustivité à la somme des N estimés dans chacune des strates des variables présentant effectivement une hétérogénéité de capture.

Comme pour l'analyse globale, une analyse de sensibilité utilisant l'estimateur de Chao a été réalisée.

Analyse par RCP

Une analyse globale a également été réalisée pour chaque RCP spécifique du département des Bouches-du-Rhône, soit 9 RCP d'Urologie, une RCP polyvalente avec un volet urologique et 3 RCP d'hématologie. Les RCP spécifiques des départements limitrophes n'ont pas été incluses dans cette analyse, le nombre de cas étant trop faible. Les résultats présentés ici sont anonymisés.

2.3.5 Outil d'analyse

La constitution de la base de données agrégée et les analyses ont été réalisées avec le logiciel R (version R.4.3.1).

Les analyses capture-recapture par modèle log-linéaires ont été réalisées avec le package R-Capture qui permet d'estimer N et d'autres paramètres statistiques en population fermée ou ouverte, de sélectionner le modèle le plus robuste parmi les différents modèles log-linéaires proposés en tenant compte des interactions entre les sources et de contrôler l'hétérogénéité de capture par différents estimateurs [18].

3. RÉSULTATS

3.1 Cas identifiés par chacune des sources

3.1.1 Revela13

Avant la mise en œuvre de l'enquête d'exhaustivité, l'observatoire comptait 2 335 cas incidents de cancers diagnostiqués entre le 1er avril 2012 et le 31 décembre 2014. Parmi ces cas, 994 avaient été signalés par des médecins spécialistes, 1 330 à partir des extractions RCP et 11 à partir de CR d'ACP. Plus de 4 000 signalements issus des extractions RCP, tous cancers confondus, étaient investigués. Le Tableau 1 présente la répartition des signalements et des cas inclus (doublons inter sources compris) par année de diagnostic pour les cancers urologiques et pour les leucémies aiguës.

I TABLEAU 1 I

Nombre et traitements des signalements issus des médecins spécialistes et des extractions de RCP par année de diagnostic et par cancer, Bouches-du-Rhône, avril 2012 –décembre 2014

Sources de signalement	Année	Cancers urologiques		Leucémies aiguës	
		Signalements	Cas inclus	Signalements	Cas inclus
Signalements directs					
Médecins spécialistes	2012	ND	204	ND	50
	2013	ND	358	ND	43
	2014	ND	330	ND	9
Signalements indirects					
Réseau régional OncoPaca-Corse	2012	536	191	72	8
	2013	959	259	243	19
	2014	1023	367	294	18
3C AP-HM	2012	142	69	8	0
	2013	195	71	24	5
	2014	170	102	16	11
3C IPC	2012-2013	113	27	154	59
	2014	142	50	95	46
3C Vaucluse	2012-2013	52	26	9	4
Réseau régional OoncoLR 3C public - privé de Nîmes	2012-2014	50	22	3	2
CR ACP	2012-2014	ND	11	-	-
Total		3 382	2 087	918	274

Parmi les 2 335 cas incidents de cancers enregistrés dans l'observatoire, 1 393 cas étaient des tumeurs urothéliales, 690 des cancers du rein et 252 des cas de leucémies aiguës. Le

Tableau 2 présente la répartition des cas inclus dans l'observatoire par RCP primo-déclarante (un cas pouvant être signalé par plusieurs RCP) avant la mise en œuvre de l'enquête d'exhaustivité. À noter que 7 cas de tumeurs urothéliales et 4 cas de cancers du rein n'avaient pas été identifiés à partir d'un passage en RCP mais à partir d'une extraction de CR anatomopathologique ciblé sur un urologue qui n'était pas encore rattaché à une RCP spécifique.

I TABLEAU 2 I

Nombre de cas de tumeurs urothéliales, de cancers du rein et de leucémies aiguës inclus par RCP, Bouches-du-Rhône, avril 2012 - décembre 2014

Zones	Intitulé de RCP	Tumeurs urothéliales	Cancers du rein	Leucémies aiguës
Marseille	RCP Urologie Prado-Louvain	385	131	-
	RCP Urologie Castellane	294	86	-
	RCP Urologie Nord	124	98	-
	RCP Urologie Conception	86	70	-
	RCP Hématologie Conception	-	-	43
	RCP Urologie IPC	51	61	-
	RCP Hématologie IPC	-	-	153
	RCP Urologie HP Beauregard - RCP Oncologie Vert Coteau	44	8	-
	RCP Oncologie Bonneveine	8	1	-
	RCP Oncologie Clairval	5	5	-
	Autres RCP non spécifiques	8	1	-
Aix-en-Provence	RCP Urologie Polyclinique du Parc Rambot	109	62	-
	RCP Urologie Axiom	31	22	-
	RCP Urologie CH du Pays d'Aix	50	39	-
	RCP Hématologie CH du Pays d'Aix	-	-	33
	Autres RCP Non spécifiques	-	1	2
Salon de Provence	RCP Urologie Vignoli	48	21	-
	RCP Polyvalente-Oncologie CH Salon	0	1	6
Marignane	RCP Oncologie ROLEB	65	33	-
Istres	RCP Oncologie Istres	4	3	-
Arles - Avignon	RCP Urologie Vaucluse	65	26	-
	RCP Hématologie CH Avignon	-	-	13
Nîmes	RCP Urologie Public Nîmes	4	3	-
	RCP Hématologie CHRU Nîmes	-	-	2
	RCP Urologie Privé Nîmes	5	10	-
Autres RCP hors département des Bouches-du-Rhône		0	4	-
Total		1 386	686	252

3.1.2 PMSI

Parmi les 9 100 signalements adressés par les vingt-sept DIM contactés, 4 315 signalements (doublons inter-établissements compris) ont été exclus pour les raisons suivantes : diagnostic antérieur au 2 avril 2012 ; résidence hors du département ; cas signalés par une RCP et exclus. Parmi les signalements retenus, 2 451 étaient déjà inclus dans la base de données Revela13 et 2 379 signalements restaient à investiguer (Tableau 3).

I TABLEAU 3 I

Nombre et traitement des signalements reçus par DIM sollicité, tous cancers confondus, Bouches-du-Rhône, avril 2012 –décembre 2014

PMSI	Signalements	Déjà inclus	Exclus	À rechercher
DIM AP-HM	2044	374	998	672
DIM Hôpital Saint-Joseph	1351	271	927	153
DIM Polyclinique du Parc Rambot	426	165	117	144
DIM Institut Paoli-Calmette	735	245	341	149
DIM Hôpital Européen	1022	294	558	170
DIM Clinique de Marignane	337	105	139	93
DIM Clinique Rhône-Durance	299	60	137	102
DIM Centre hospitalier Joseph Imbert	46	8	7	31
DIM Groupe Sainte-Marguerite**	387	174	123	90
DIM Hôpital privé Résidence du Parc	217	28	98	91
DIM Clinique Jeanne d'Arc	94	23	15	56
DIM Clinique Axium	175	64	81	30
DIM Centre hospitalier du Pays d'Aix	272	119	71	82
DIM Clinique Bonneveine	126	27	64	35
DIM Clinique Bouchard	117	33	6	78
DIM CHRU Nîmes	63	16	27	20
DIM Hôpital privé de la Casamance	428	157	208	63
DIM Centre hospitalier de Martigues	117	29	34	54
DIM Clinique de l'Étang de l'Olivier	128	24	66	38
DIM Clinique de Martigues	142	30	39	73
DIM Clinique Vignoli	287	62	172	53
DIM Centre hospitalier Edmond Garcin	32	8	5	19
DIM Centre hospitalier de Salon de Provence	61	22	8	31
DIM Clinique Chantecler	166	91	40	35
DIM Hôpital privé Clairval	25	6	13	6
DIM Centre hospitalier Henri Duffaut	31	11	9	11
DIM Polyclinique Kenval	17	5	12	0
Total	9 145	3 451	4 315	2 379

** Clinique Vert Coteau, Clinique de la Ciotat, Hôpital Beauregard

Après retour au médecin référent, 402 cas (doublons inter-établissements inclus) ont été confirmés pour les trois cancers étudiés dont 245 cas de tumeurs urothéliales, 87 cas de cancer du rein et 70 cas de leucémies aiguës. Quarante-huit cas n'ont pu être validés soit parce que les 2 chirurgiens-urologues prescripteurs ont refusé de participer à l'étude (18 cas), soit parce que les dossiers des patients n'ont pu être retrouvés.

L'ensemble des DIM sollicités ont conduit à l'inclusion d'au moins un cas, à l'exception du DIM de la polyclinique Kenval du Gard pour lequel tous les cas transmis étaient déjà inclus (Tableau 4).

I TABLEAU 4 I

Nombre de cas de cancers urinaires et de leucémies aiguës inclus pour chaque DIM sollicité, Bouches-du-Rhône, avril 2012 - décembre 2014

PMSI	Cas inclus		Cas exclus	Cas non retrouvés
	Cancers urinaires	Leucémies aiguës		
DIM AP-HM	83	12	575	2
DIM Hôpital Saint-Joseph	27	4	115	7
DIM Polyclinique du Parc Rambot	29	-	115	0
DIM Institut Paoli-Calmette	22	35	92	0
DIM Hôpital Européen	20	-	150	0
DIM Clinique de Marignane	20	-	65	8
DIM Clinique Rhône-Durance	17	-	84	1
DIM Centre hospitalier Joseph Imbert	7	1	23	0
DIM Groupe Sainte-Marguerite**	13	1	75	1
DIM Hôpital Privé Résidence du Parc	11	1	54	21
DIM Clinique Jeanne d'Arc	11	-	44	1
DIM Clinique Axium	10	-	20	0
DIM Centre hospitalier du Pays d'Aix	8	9	65	0
DIM Clinique Bonneveine	8	1	26	0
DIM Clinique Bouchard	8	-	70	0
DIM CHRU Nîmes	8	-	12	0
DIM Hôpital privé de la Casamance	6	-	54	3
DIM Centre hospitalier de Martigues	5	1	48	0
DIM Clinique de l'Étang de l'Olivier	5	-	33	0
DIM Clinique de Martigues	5	-	67	1
DIM Clinique Vignoli	4	-	49	0
DIM Centre hospitalier Edmond Garcin	2	-	15	0
DIM Centre hospitalier de Salon de Provence	2	-	27	2
DIM Clinique Chantecler	1	-	34	0
DIM Hôpital privé Clairval	0	2	4	0
DIM Centre hospitalier Henri Duffaut	0	1	9	1
Total	332	70	1 925	48

** Clinique Vert Coteau, Clinique de la Ciotat, Hôpital Privé Beauregard

Sur les 70 cas de leucémies aiguës retrouvés par les DIM, 6 étaient communs à deux DIM et six avaient un diagnostic postérieur à 2014. Ces derniers ont été inclus dans la base de données Revela13 mais n'ont pas été pris en compte dans l'analyse d'exhaustivité qui portait sur la période 2013- 2014. Au total, 56 cas supplémentaires de leucémies aiguës ont été inclus dans la base de données Revela13.

De même, parmi les 245 tumeurs urothéliales incluses sur signalement d'un DIM, 7 patients avaient un diagnostic de tumeur invasive postérieur à décembre 2014. Ces cas n'ont été pas retenus pour l'étude d'exhaustivité.

3.1.3 Laboratoires ACP

L'observatoire Revela13 a reçu 4 390 signalements de patients des 13 laboratoires participants. Un grand nombre de ces signalements ne répondaient pas aux critères d'inclusion (hors département ou hors diagnostic) ou avaient déjà été exclus dans les signalements de RCP.

Parmi les signalements reçus (doublons inclus), 1 872 correspondaient à des cas déjà enregistrés par l'observatoire dont 1 286 tumeurs urothéliales et 586 cancers du rein. Il restait 1 853 signalements à vérifier dont 1 430 tumeurs urothéliales et 423 cancers du rein (Tableau 5).

I TABLEAU 5 I

Nombre et traitement des signalements reçus par laboratoire sollicité, Bouches-du-Rhône, avril 2012 –décembre 2014

Laboratoires d'anatomopathologie	Signalements	Déjà inclus	Exclus	À rechercher
ACP Hôpital de la Timone (AP-HM)	841	190	63	588
ACP Hôpital Nord (AP-HM)	521	190	13	318
SCP Drs Chollat-Namy Calvet Patte (Marseille)	501	281	59	161
Massalia Pathologie (Marseille)	759	350	293	116
SCM Drs Aubert-Severy-Henric (Martigues)	663	251	130	282
ACP Hôpital Saint-Joseph	301	229	4	68
ACP Institut Paoli Calmette	286	67	24	195
ACP Centre hospitalier du Pays d'Aix	103	76	14	13
HISTOSUD (Le Pontet)	76	52	10	14
MEDIPATH (Eguilles)	228	148	31	49
SCP Dr Niel - Bourelly (Martigues)	16	10	1	5
SUDPATH (La Ciotat)	66	14	21	31
SCP Dr Masset (Arles)	29	14	2	13
Total	4 390	1 872	665	1 853

Le retour au médecin prescripteur a permis au final d'inclure 307 cas dont 231 tumeurs urothéliales et 76 cancers du rein. Trente-huit signalements n'ont pas pu être vérifiés du fait du refus de deux urologues de participer à l'étude (20 patients) ou de dossiers non retrouvés au cabinet médical référent.

Parmi les 231 tumeurs urothéliales issues des données ACP, 12 patients avaient un diagnostic de tumeur invasive postérieur à décembre 2014, ces signalements n'ont été pas retenus dans l'analyse d'exhaustivité.

L'ensemble des laboratoires sollicités a permis de retrouver des cas incidents non enregistrés par l'observatoire Revela13. Le nombre de nouveaux cas inclus par type de cancer pour chaque laboratoire est présenté dans le Tableau 6.

I TABLEAU 6 I

Nombre de cas de cancers inclus par laboratoire sollicité et par type de cancer, Bouches-du-Rhône, avril 2012 - décembre 2014

Laboratoires d'anatomopathologie	Cas inclus		Cas exclus	Cas non retrouvés
	Tumeurs urothéliales	Cancers du rein		
ACP – Hôpital de la Timone (AP-HM)	44	32	511	1
ACP – Hôpital Nord (AP-HM)	14	3	301	0
SCP Drs Chollat-Namy Calvet Patte (Marseille)	23	4	100	34
Massalia Pathologie (Marseille)	42	0	73	1
SCM Drs Aubert-Severy-Henric (Martigues)	39	3	239	1
ACP Hôpital Saint-Joseph	17	10	41	0
ACP Institut Paoli-Calmette	12	11	172	0
ACP Centre hospitalier du Pays d'Aix	4	4	5	0
HISTOSUD (Le Pontet)	8	2	4	0
MEDIPATH (Eguilles)	12	7	29	1
SCP Dr Niel - Bourelly (Martigues)	4	0	1	0
SUDPATH (La Ciotat)	2	0	29	0
SCP Dr Masset (Arles)	10	0	3	0
Total	231	76	1 508	38

Au final, sur l'ensemble de la période considérée, les DIM ont permis d'inclure 2 510 cas, les laboratoires 2 067 alors que l'observatoire en a enregistré 2 335 sur la même période (Tableau 7). En croisant les cas non retrouvés pour les DIM et les ACP, 64 signalements n'ont pu être vérifiés : 39 signalements de tumeurs urothéliales, 23 de cancers du rein et 2 de leucémies aiguës. Parmi ces signalements, 26 correspondaient à des patients dont le médecin à l'origine du diagnostic avait refusé de participer à l'observatoire Revela13. Les autres signalements n'ont pu être vérifiés faute de retrouver les dossiers médicaux correspondants.

I TABLEAU 7 I

Nombre de cas de cancers urinaires et de leucémies aiguës inclus pour chaque source, Bouches-du-Rhône, avril 2012 - décembre 2014

	Tumeurs urothéliales	Rein	Leucémies aiguës	Total
Nombre de cas inclus Revela13	1393	690	252	2 335
Nombre de cas inclus PMSI	1533	703	274	2 510
Nombre de cas inclus ACP	1436	631		2 067
Nombre de cas non retrouvés (PMSI + ACP)	39	23	2	64
dont refus	20	6	0	26

Compte tenu, de la mise en route progressive de l'observatoire Revela13 en 2012 et du recueil incomplet démarré seulement à partir du second trimestre, l'année 2012 n'a pas été conservée pour les analyses de l'étude capture-recapture. Le Tableau 8 présente les données utilisées dans l'étude capture-recapture.

I TABLEAU 8 I

Nombre de cas de cancers urinaires et de leucémies aiguës inclus pour chaque source, Bouches-du-Rhône, 2013-2014

	Tumeurs urothéliales	Rein	Leucémies aiguës	Total
Nombre de cas inclus Revela13	1 047	505	177	1 729
Nombre de cas inclus PMSI	1 141	523	195	1 859
Nombre de cas inclus ACP	1 082	487		1 569

3.2 Exhaustivité pour les leucémies aiguës

3.2.1 Description des données par sources

Pour les leucémies aiguës, l'analyse capture-recapture a été réalisée avec 2 sources en raison de l'impossibilité de recueillir les myélogrammes spécifiques aux leucémies aiguës.

La distribution des cas par sexe, classe d'âge, zone de résidence et type de leucémie était similaire dans les sources Revela13 et PMSI (Tableau 9). Le sexe ratio et la distribution par âge correspondaient à ce qui est décrit dans la littérature.

I TABLEAU 9 I

Distribution des cas déclarés de leucémies aiguës dans les deux sources selon l'âge, le sexe, la zone de résidence et le type, Bouches-du-Rhône, 2013-2014

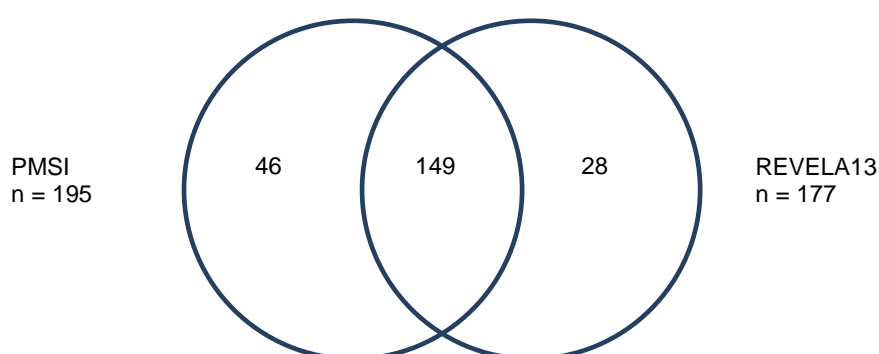
Caractéristiques	Revela13	PMSI	Significativité
Sexe			
Homme	95	103	0,9
Femme	82	92	
Classe d'âge			
Moins de 60 ans	65	66	0,8
60-79 ans	75	87	
80 ans et plus	37	42	
Secteur de résidence			
Marseille	76	84	1
Hors Marseille	101	111	
Type de leucémies aiguës			
LAM	152	166	0,9
LAL	23	26	

3.2.2 Analyse globale

Avant l'enquête d'exhaustivité, l'observatoire Revela13 comptait 177 cas de leucémies aiguës diagnostiqués sur la période 2013-2014. Durant la même période, 195 cas ont été enregistrés par le PMSI dont 149 cas communs aux deux sources (Figure 1). Au total, 223 cas ont été capturés par les deux sources sur la période retenue dont 46 cas uniquement à partir du PMSI (21 %).

I FIGURE 1 I

Répartition des leucémies aiguës au sein des deux sources, Bouches-du-Rhône, 2013-2014



L'application de la méthode capture-recapture à deux sources retrouve un nombre total de cas estimé de 232 [IC95% : 225-239].

Le taux d'exhaustivité de Revela13 est de 76,3 % [IC95% : 74,0 % - 78,8 %] et celui de la source PMSI de 84,0 % [IC95% : 81,5 % - 86,8 %].

L'exhaustivité globale pour les deux sources est de 96,1 % [IC95% : 93,3 % - 99,1 %].

3.2.3 Analyse stratifiée

Des analyses stratifiées sur différentes variables d'hétérogénéité ont été réalisées (Tableau 10).

L'âge était à l'origine d'une hétérogénéité de capture pour les deux sources. La probabilité de capture diminuait en effet avec les classes d'âges les plus âgées, de 60 à 79 ans puis pour les plus de 80 ans. Les personnes âgées de plus de 80 ans étaient également moins bien identifiées par la source PMSI.

Les autres variables étudiées ne montraient pas d'hétérogénéité de capture que ce soit pour la source Revela13 ou pour la source PMSI.

La somme des estimations de N pour chaque classe d'âge a conduit à un nombre total de cas estimé de 236, à peine plus élevé que celui estimé par le modèle global. Le taux d'exhaustivité de Revela13 a diminué légèrement de 76,3 % à 75,1 %.

Après prise en compte de l'hétérogénéité de capture liée à l'âge, l'exhaustivité des deux sources pour les leucémies aiguës diminue de 96,1 % à 94,5 %.

I TABLEAU 10 I

Estimations du nombre de cas de leucémies aiguës et du taux d'exhaustivité des sources stratifiées sur les variables d'hétérogénéité, Bouches-du-Rhône, 2013-2014

	N observé	N estimé (IC 95%)	Exhaustivité des sources (%) (IC 95%)	
			Revela13	PMSI
Année de diagnostic				
2013	123	127 (122 - 131)	78,9 (76,2 - 81,8)	86,8 (83,8 - 90,0)
2014	100	106 (100 - 112)	72,8 (68,7 - 77,3)	80,3 (75,8 - 85,4)
Sexe				
Homme	116	120 (115 - 124)	79,5 (76,6 - 82,6)	86,2 (83,1 - 89,6)
Femme	107	113 (107 - 119)	72,6 (68,7 - 86,2)	81,4 (77,1 - 86,2)
Classe d'âge				
Moins de 60 ans	74	75 (72 - 77)	86,3 (83,5 - 89,4)	87,6 (84,8 - 90,8)
De 60 à 79 ans	95	98 (94 - 102)	76,6 (73,5 - 79,9)	88,9 (85,3 - 92,8)
Plus de 80 ans	54	62 (53 - 71)	59,3 (51,8 - 69,2)	67,3 (58,8 - 78,6)
Secteur de résidence				
Marseille	99	105 (99 - 111)	72,6 (68,6 - 77,1)	80,2 (75,8 - 85,2)
Hors Marseille	124	126 (123 - 132)	79,2 (76,5 - 82,1)	87,1 (84,1 - 90,2)
Type de leucémies aiguës				
LAM	190	198 (191 - 204,2)	77,0 (74,4 - 79,7)	84,0 (81,3 - 87,0)
LAL	30	32 (29 - 35)	72,8 (66,3 - 80,7)	82,3 (74,9 - 91,2)

3.2.4 Analyse par RCP d'hématologie

L'analyse globale réalisée pour les 3 RCP spécifiques d'hématologie du département a mis en évidence des taux d'exhaustivité différents pour ce qui concerne la source Revela13, alors que la probabilité de capture d'un cas de leucémie aiguë par le PMSI était similaire pour ces 3 RCP (Tableau 11). L'exhaustivité des deux sources varie de 97,8 % à 98,8 % selon les RCP.

I TABLEAU 11 I

Estimation de l'exhaustivité de chaque source pour les leucémies aiguës en fonction des RCP d'hématologie, Bouches-du-Rhône, 2013-2014

RCP Hématologie	Revela13 (%)	PMSI (%)	2 sources (%)
RCP Hémato - A	82,5	89,1	98,0
RCP Hémato - B	78,9	91,5	97,8
RCP Hémato - C	86,7	91,6	98,8

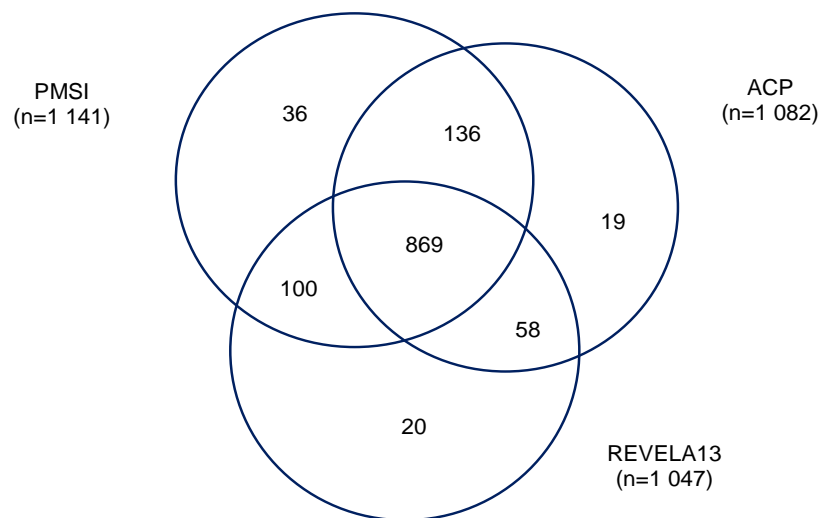
3.3 Exhaustivité pour les tumeurs urothéliales

3.3.1 Description des données par source

Avant l'enquête d'exhaustivité, la base de données Revela13 comptait 1 047 cas de tumeurs urothéliales. Durant la même période, 1 141 cas ont été enregistrés par les DIM et 1 082 par les laboratoires ACP. Un total de 1 238 cas incidents a été notifié dont 869 cas communs aux trois sources (Figure 2).

I FIGURE 2 I

Répartition des tumeurs urothéliales au sein des trois sources, Bouches-du-Rhône, 2013-2014



La distribution des cas par sexe, classe d'âge, secteur de résidence et stade tumoral était similaire pour les sources Revela13, PMSI et ACP (Tableau 12). Le sexe ratio et la distribution par âge correspondaient à ce qui est décrit dans la littérature.

I TABLEAU 12 I

Distribution des cas déclarés de tumeurs urothéliales dans les trois sources selon l'âge, le sexe, la zone de résidence, la localisation et le stade, Bouches-du-Rhône, 2013-2014

Caractéristiques	Revela13	PMSI	ACP	Significativité
Sexe				
Homme	827	903	849	0,9
Femme	220	238	233	
Classe d'âge				
Moins de 65 ans	263	279	252	0,8
65-84 ans	643	695	667	
85 ans et plus	141	167	163	
Secteur de résidence				
Marseille	510	543	527	0,8
Hors Marseille	537	598	555	
Localisation tumorale				
Vessie	948	1 033	978	0,9
Autres localisations	99	108	104	
Stade				
TVNIM	584	655	612	0,8
TVIM	463	486	470	

3.3.2 Dépendance entre sources - analyse globale

La recherche de la dépendance entre deux sources peut être effectuée au sein la 3^{ème} source par le calcul d'un rapport de cote (OR) selon la méthode de Wittes.

Une dépendance positive a été retrouvée par chacune des sources prises deux à deux, indiquant une probabilité plus élevée pour chaque patient d'être présent dans une source lorsqu'il a été identifié dans les deux autres sources (Tableau 13). Cette dépendance positive peut conduire à une sous-estimation du nombre total de cas et donc à une surestimation de l'exhaustivité.

I TABLEAU 13 I

Estimation de la dépendance des sources de notification dans la détection des cas de tumeurs urothéliales, Bouches-du-Rhône, 2013-2014

	ACP		OR	PMSI		OR
	Présents	Absents	[IC 95%]	Présents	Absents	[IC 95%]
Revela13			2,3 [1,5 - 3,5]			2,1 [1,1 - 3,7]
Présents	869	100		869	58	
Absents	136	36		136	19	
PMSI			2,9 [1,6 - 5,2]			
Présents	869	100				
Absents	58	20				

Selon les différents modèles de l'analyse globale, le nombre total de cas estimé variait entre 1 247 et 1 253. Le modèle présentant les plus petits AIC et BIC dans l'analyse globale était celui sans interaction entre les sources avec un nombre de cas total estimé à 1 247 [IC 95% : 1 239 – 1 255] (Tableau 14).

I TABLEAU 14 I

Estimation du nombre de cas total de tumeurs urothéliales (N) selon les différents modèles log-linéaires, Bouches-du-Rhône, 2013-2014

Modèle	N	ET	Déviante	ddl	AIC	BIC
Revela13 ACP PMSI	1 247,2	4,2	1,159	2	53,921	79,527
Revela13 ACP PMSI Revela13*ACP	1 247,4	4,3	1,072	1	55,436	86,163
Revela13 ACP PMSI Revela13*PMSI	1 246,7	4,1	0,674	1	55,436	86,163
Revela13 ACP PMSI ACP*PMSI	1 248,4	4,7	0,083	1	54,845	85,573
Revela13 ACP PMSI Revela13*ACP Revela13*PMSI	1 246,7	4,1	0	0	56,762	92,611
Revela13 ACP PMSI Revela13*ACP ACP*PMSI	1 248,4	4,7	0	0	56,762	92,611
Revela13 ACP PMSI Revela13*PMSI ACP*PMSI	1 247,9	4,8	0	0	56,762	92,611
Revela13 ACP PMSI Revela13*ACP Revela13* PMSI ACP*PMSI	1 253,1	7,3	-	0	56,762	92,611

N : nombre total de cas estimé ; ET : écart-type ; ddl : degrés de liberté, AIC : critère d'Akaike ; BIC : critère d'information bayésien

Le taux d'exhaustivité de Revela13 pour les tumeurs urothéliales était de 83,9 % [IC95 % : 83,4 %-84,5 %] et était significativement inférieur à celui des sources PMSI et ACP (Tableau 15).

I TABLEAU 15 I

Estimation de l'exhaustivité et intervalle de confiance à 95% de chaque source pour les tumeurs urothéliales, Bouches-du-Rhône, 2013-2014

Source	Exhaustivité (%)	IC95%
Revela13	83,9	83,4 - 84,5
PMSI	91,5	90,9 - 92,1
ACP	86,7	86,2 - 87,3
3 sources	99,3	98,6 – 99,9

Les modèles de régression log-linéaires réalisés avec l'estimateur de Chao montraient des résultats très proches de ceux observés avec l'estimateur de Poisson que ce soit en termes de nombre de cas ou d'exhaustivité (Annexe 9).

3.3.3 Analyse stratifiée

Différentes variables potentielles d'hétérogénéité ont été étudiées par des analyses stratifiées suivant les modèles log-linéaires sélectionnés (Tableau 16).

I TABLEAU 16 I

Estimations du nombre de cas de tumeurs urothéliales et du taux d'exhaustivité des sources stratifiées sur les variables d'hétérogénéité, Bouches-du-Rhône, 2013-2014

	N observé	N estimé (95% IC)	Exhaustivité des sources (%) (IC95%)		
			Revela13	PMSI	ACP
Année de diagnostic					
2013	627	632 (626 - 637)	82,5 (81,7 - 83,2)	91,0 (90,2 - 92,0)	86,4 (85,7 - 87,2)
2014	611	616 (610 - 622)	85,4 (84,6 - 86,3)	91,9 (91,1 - 92,8)	87,0 (86,2 - 87,9)
Sexe					
Homme	973	982 (974 - 991)	84,2 (83,3 - 84,9)	91,9 (91,1 - 92,8)	86,4 (85,7 - 87,2)
Femme	265	266 (264 - 268)	82,8 (82,2 - 83,4)	89,5 (88,9 - 90,2)	87,7 (87,0 - 88,3)
Classe d'âge					
Moins de 65 ans	293	294 (292 - 297)	89,3 (88,5 - 90,2)	94,8 (93,8 - 95,7)	85,6 (84,8 - 86,4)
65 - 84 ans	765	771 (764 - 777)	83,4 (82,7 - 84,2)	90,1 (89,4 - 90,9)	86,5 (85,8 - 87,3)
85 ans et plus	180	182 (178 - 186)	77,4 (75,6 - 79,3)	91,7 (89,6 - 93,9)	89,5 (87,4 - 91,7)
Secteur de résidence					
Marseille	583	584 (582 - 586)	87,3 (87,0 - 87,7)	93,0 (92,7 - 93,3)	90,3 (89,9 - 90,6)
Hors Marseille	655	667 (656 - 678)	80,5 (79,1 - 81,8)	89,6 (88,1 - 91,1)	83,1 (81,8 - 84,6)
Stade					
TVNIM	706	711 (705 - 716)	82,2 (81,5 - 82,8)	91,9 (91,2 - 92,6)	85,8 (85,2 - 86,5)
TVIM	532	537 (531 - 543)	86,2 (85,2 - 87,2)	90,9 (89,8 - 92,0)	87,9 (86,8 - 88,9)

Année de diagnostic

L'année de diagnostic était à l'origine d'une hétérogénéité de capture avec une probabilité plus élevée de capture par la source Revela13 pour les cas diagnostiqués en 2014 par rapport à ceux diagnostiqués en 2013, expliquant une meilleure exhaustivité de recueil cette année-là. L'hétérogénéité de capture liée à l'année de diagnostic n'a été pas retrouvée dans les deux autres sources.

Sexe

Les femmes atteintes d'une tumeur urothéliale étaient significativement moins bien détectées par la source DIM que les hommes, une tendance similaire non significative a été observée pour la source Revela13.

Âge

Le taux d'exhaustivité de la source Revela13 diminuait progressivement et significativement avec l'âge passant de de 89,3 % chez les moins de 65 ans à 77,4 % pour les plus de 85 ans. Pour la source PMSI, l'exhaustivité était supérieure à 90% quel que soit l'âge et un peu plus élevée pour les moins de 65 ans. A contrario, pour la source ACP l'exhaustivité était meilleure pour les cas les plus âgés.

Secteur de résidence

La probabilité de capture était également significativement plus grande chez les cas résidant à Marseille comparé à ceux résidant hors Marseille quelle que soit la source de donnée considérée.

Stade tumoral

Enfin, les TVIM (stade > pT1) étaient également mieux identifiées par la source Revela13 et par la source ACP que les TVNIM.

L'âge, le secteur de résidence, le type de tumeurs urothéliales et, dans une moindre mesure le sexe, étaient à l'origine d'une hétérogénéité de capture quelle que soit la source considérée. Une analyse stratifiée par zone de résidence et classe d'âge a été réalisée pour chaque sexe. Les résultats sont présentés en détails en Annexe 10. L'exhaustivité de la source Revela13 était meilleure chez les patients les plus jeunes (moins de 65 ans) résidant à Marseille indépendamment du sexe (Tableaux 17 et 18).

Chez les hommes, pour les sources Revela13 et PMSI, la probabilité de capture diminuait progressivement et significativement avec l'âge chez ceux résidant hors de Marseille (Tableau 17). La probabilité de capture la plus faible est ainsi observé chez les hommes âgés de plus de 85 ans et résidant en dehors de Marseille (72,7 %).

I TABLEAU 17 I

Estimation de l'exhaustivité de chaque source pour les tumeurs urothéliales chez les hommes par classe d'âge et secteur de résidence, Bouches-du-Rhône, 2013-2014

	Exhaustivité (%) Marseille			Exhaustivité (%) Hors Marseille		
	Moins de 65 ans	65 à 84 ans	85 ans et plus	Moins de 65 ans	65 à 84 ans	85 ans et plus
Revela13	88,5	86,9	84,9	88,7	80,5	72,7
PMSI	94,4	93,7	96,2	94,9	87,7	84,5
ACP	91,5	90,0	92,4	81,7	82,6	87,3

Chez les femmes, les probabilités de capture ne variaient pas de façon homogène au sein d'une même source entre les différentes classes d'âge, et entre les sources pour chacune des classes d'âge considérées, sauf chez les femmes résidant hors Marseille, dont l'exhaustivité diminuait avec l'âge pour la source Revela13 (Tableau 18). Comme chez les hommes, la plus faible probabilité de capture par la source Revela13 est ainsi observée chez les femmes âgées de plus de 85 ans résidant hors de Marseille (62 %).

I TABLEAU 18 I

Estimation de l'exhaustivité de chaque source pour les tumeurs urothéliales chez les femmes par classe d'âge et secteur de résidence, Bouches-du-Rhône, 2013-2014

	Exhaustivité (%) Marseille			Exhaustivité (%) Hors Marseille		
	Moins de 65 ans	65 à 84 ans	85 ans et plus	Moins de 65 ans	65 à 84 ans	85 ans et plus
Revela13	89,5	87,5	88,0	92,5	75,9	62,3
PMSI	94,7	84,7	92,0	92,5	85,5	95,1
ACP	78,9	91,6	96,0	89,0	82,8	88,5

La somme des estimations de N pour chaque strate (classe d'âge*sexe*secteur de résidence) a conduit à un nombre total de cas estimé de 1 253 (Annexe 12), à peine plus élevé que celui estimé par le modèle global. L'exhaustivité de la source Revela13 passe ainsi de 83,9 % à 83,6 %.

Après prise en compte de l'hétérogénéité de capture liée à l'âge, le sexe et le secteur de résidence, l'exhaustivité des trois sources pour les tumeurs urothéliales, diminue de 99,3 % à 98,8 %.

3.3.4 Analyse par RCP d'urologie

Une analyse globale a été réalisée pour chacune des 9 RCP spécifiques d'urologie et d'une RCP polyvalente du département (Tableau 19). L'exhaustivité pour les tumeurs urothéliales était très variable d'une RCP à l'autre quelle que soit la source considérée. Par ailleurs pour une même RCP, l'exhaustivité variait également selon la source de données considérée. L'exhaustivité des trois sources variait de 93 % à 100 % selon les RCP.

I TABLEAU 19 I

Estimation de l'exhaustivité de chaque source pour les tumeurs urothéliales en fonction des RCP d'urologie ou polyvalente, Bouches-du-Rhône, 2013-2014

RCP Urologie	Revela13 (%)	PMSI (%)	ACP (%)	3 Sources (%)
RCP Urologie - D	94,5	96,1	89,3	99,9
RCP Urologie - E	92,7	84,8	87,5	99,3
RCP Urologie - F	89,5	93,3	90,5	99,9
RCP Urologie - G	82,3	97,1	97,1	100,0
RCP Urologie - H	84,9	87,9	75,9	95,9
RCP Urologie - I	87,2	82,8	93,9	98,4
RCP Polyvalente - J	84,3	97,3	90,8	99,9
RCP Urologie - K	73,4	88,2	93,7	99,1
RCP Urologie - L	94,1	96,0	93,1	100,0
RCP Urologie - M	82,0	79,8	75,4	93,1

3.4 Exhaustivité des cancers du rein

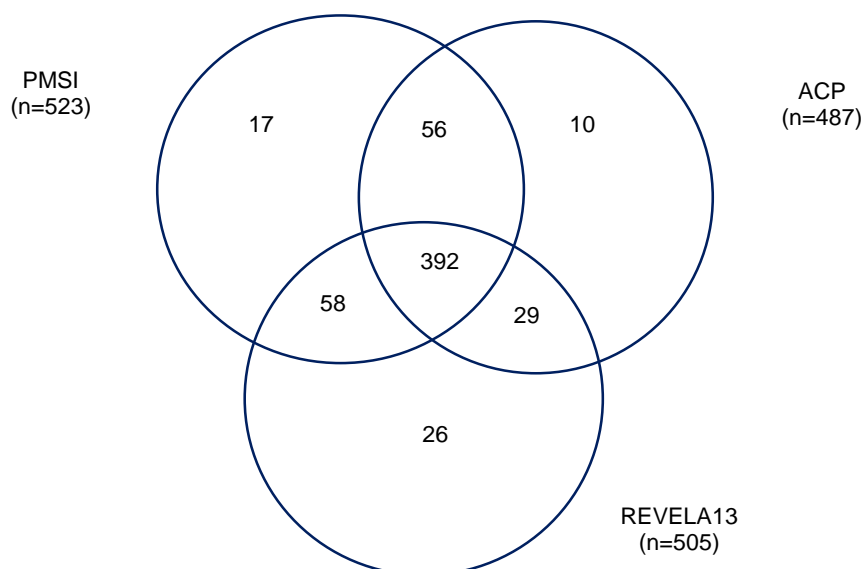
3.4.1 Description des données par source

Avant l'enquête d'exhaustivité, l'observatoire Revela13 comptait 514 cas de tumeurs du rein dont 505 avec un diagnostic histologique et 9 avec un diagnostic par imagerie, ces derniers n'ont pas été retenus dans l'analyse, car ils ne pouvaient pas être détectés par la source ACP.

Durant la même période, 525 cas ont été enregistrés par la source PMSI (dont 2 avec un diagnostic par imagerie) et 487 par les laboratoires ACP. Un total de 588 cas incidents avec une histologie confirmée a été notifié, dont 392 communs aux trois sources (Figure 3).

I FIGURE 3 I

Répartition des tumeurs du rein au sein des trois sources, Bouches-du-Rhône, 2013 – 2014



La distribution des cas par sexe, classe d'âge, zone de résidence et stade tumoral était similaire dans les sources Revela13, PMSI et ACP (Tableau 20). Le sexe ratio et la distribution par âge correspondaient à ce qui est décrit dans la littérature.

I TABLEAU 20 I

Distribution des cas déclarés de cancers du rein dans les trois sources selon l'âge, le sexe, la zone de résidence, la localisation et le stade, Bouches-du-Rhône, 2013-2014

Caractéristiques	Revela13	PMSI	ACP	Significativité
Année de diagnostic				
2013	238	252	235	0,9
2014	267	271	252	
Sexe				
Homme	342	355	330	1,0
Femme	163	168	157	
Classe d'âge				
Moins de 55 ans	115	118	107	1,0
55-74 ans	271	286	268	
75 ans et plus	119	119	119	
Zone de résidence				
Marseille	197	201	193	0,9
Hors Marseille	308	322	294	
Stade				
pT1	277	288	275	0,7
>pT1	176	179	157	

3.4.2 Dépendance entre sources - analyse globale

Une dépendance positive a été retrouvée par chacune des sources prises 2 à 2, indiquant une probabilité plus élevée pour chaque patient d'être présent dans une source lorsqu'il était identifié dans les deux autres sources. Cette dépendance n'était cependant pas homogène entre les trois sources puisqu'elle était plus élevée pour les sources PMSI et ACP (Tableau 21).

I TABLEAU 21 I

Dépendance des sources de notification dans la détection des cas de tumeurs du rein, Bouches-du-Rhône, 2013 - 2014

	ACP		OR	PMSI		OR
	Présents	Absents	[IC 95%]	Présents	Absents	[IC 95%]
Revela13			2,0 [1,0 - 3,9]			2,4 [1,0 - 5,4]
Présents	392	58		392	29	
Absents	56	17		56	10	
PMSI			6,0 [3,2 - 11,4]			
Présents	392	58				
Absents	29	26				

Selon les différents modèles de l'analyse globale, le nombre total de cas estimé variait entre 597 et 606. Le modèle présentant les plus petits AIC et BIC dans l'analyse globale était celui avec interaction entre les sources PMSI et ACP estimant le nombre de cas total à 600 [IC95% : 589-611] (Tableau 22). Cette interaction correspondait à la dépendance entre sources la plus importante.

I TABLEAU 22 I

Estimation du nombre de cas total de cancers du rein (N) selon les différents modèles log-linéaires, Bouches-du-Rhône, 2013-2014

Modèle	N	ET	Déviance	ddl	AIC	BIC
Revela13 ACP PMSI	599,3	5,5	9,392	2	58,120	80,004
Revela13 ACP PMSI Revela13*ACP	599,1	5,4	5,236	1	55,964	82,224
Revela13 ACP PMSI Revela13*PMSI	597,4	5,0	8,085	1	58,813	85,074
Revela13 ACP PMSI ACP*PMSI	599,9	5,8	0,125	1	50,853	77,113
Revela13 ACP PMSI Revela13*ACP Revela13*PMSI	595,5	4,3	0	0	52,728	83,365
Revela13 ACP PMSI Revela13*ACP ACP*PMSI	599,8	5,8	0	0	52,728	83,365
Revela13 ACP PMSI Revela13*PMSI ACP*PMSI	600,8	6,7	0	0	52,728	83,365
Revela13 ACP PMSI Revela13*ACP Revela13*PMSI ACP*PMSI	606,4	10,5	0	0	52,728	83,365

N : nombre total de cas estimé ; ET : écart-type ; ddl : degrés de liberté, AIC : critère d'Akaike ; BIC : critère d'information bayésien

Le taux d'exhaustivité de la source Revela13 a été estimé à 84,2 % [IC95 % : 82,6 % - 85,8 %], un peu inférieur à celui de la source PMSI. L'exhaustivité de la source ACP était significativement plus basse que celle des deux autres sources (Tableau 23).

I TABLEAU 23 I

Estimation de l'exhaustivité des cancers du rein et intervalle de confiance à 95% de chaque source, Bouches-du-Rhône, 2013-2014

Source	Exhaustivité (%)	IC95%
Revela13	84,2	82,6 – 85,8
PMSI	87,1	85,5 – 88,9
ACP	81,2	79,7 – 82,7
3 sources	98,0	96,2 – 99,9

Les modèles de régression log-linéaires réalisés avec l'estimateur de Chao montraient des résultats très proches de ceux observés avec l'estimateur de Poisson que ce soit en termes de nombre de cas ou d'exhaustivité (Annexe 9).

3.4.3 Analyse stratifiée

Des analyses stratifiées selon les variables potentielles d'hétérogénéité ont été réalisées selon les modèles log-linéaires sélectionnés (Tableau 24).

Pour le cancer du rein, l'âge était un facteur d'hétérogénéité présent dans les 3 sources. Les patients âgés de plus de 75 ans avaient une probabilité plus faible d'être identifiés que les autres classes d'âge. L'exhaustivité était de 72,7 % pour la source ACP et 77,3 % pour les 2 autres sources.

Pour la source ACP, la probabilité de capture était moins élevée pour les patients ayant une tumeur de stade plus élevé (> pT1) comparée aux patients avec une tumeur de stade pT1.

I TABLEAU 24 I

Estimations du nombre de cas de cancers du rein et du taux d'exhaustivité des sources stratifiées sur les variables d'hétérogénéité, Bouches-du-Rhône, 2013-2014

	N observé	N estimé (IC95%)	Exhaustivité des sources (%) (IC95%)		
			Revela13	PMSI	ACP
Année de diagnostic					
2013	277	281 (275 - 287)	84,7 (83,0 - 86,6)	89,7 (87,9 - 91,7)	83,7 (82,0 - 85,5)
2014	311	318 (309 - 327)	83,9 (81,6 - 86,3)	85,1 (82,8 - 87,6)	79,2 (77,0 - 81,6)
Sexe					
Homme	397	402 (396 - 409)	84,9 (83,4 - 86,5)	88,2 (86,6 - 89,8)	81,9 (80,5 - 83,5)
Femme	191	198 (188 - 209)	82,2 (77,9 - 86,8)	84,7 (80,3 - 89,6)	79,1 (75,0 - 83,7)
Classe d'âge					
moins de 55 ans	135	139 (132 - 145)	82,8 (79,1 - 87,1)	85,0 (81,2 - 89,3)	77,1 (73,6 - 81,0)
55 - 74 ans	311	313 (309 - 316)	86,7 (85,8 - 87,6)	91,5 (90,5 - 92,5)	85,7 (84,8 - 86,6)
75 ans et plus	142	154 (137 - 171)	77,3 (69,6 - 87,0)	77,3 (69,6 - 87,0)	72,7 (65,5 - 81,9)
Secteur de résidence					
Marseille	232	234 (230 - 238)	84,1 (82,6 - 85,6)	85,8 (84,3 - 87,3)	82,4 (81,0 - 83,8)
Hors Marseille	356	367 (355 - 380)	83,9 (81,1 - 86,9)	87,7 (84,7 - 90,8)	80,1 (77,4 - 82,9)
Stade					
pT1	316	323 (314 - 332)	85,9 (83,6 - 88,4)	89,3 (86,9 - 91,9)	85,3 (83,0 - 87,7)
>pT1	196	197 (194 - 200)	89,2 (87,8 - 90,6)	90,7 (89,3 - 92,2)	79,6 (78,3 - 80,8)

La somme des estimations de N pour chaque classe d'âge a conduit à estimer un nombre total à 606, à peine plus élevé que celui estimé par le modèle global avec une exhaustivité de la source Revela13 qui passait de 84,2 % à 83,3 %.

Après prise en compte de l'hétérogénéité de capture liée à l'âge, l'exhaustivité des trois sources pour les cancers du rein diminue de 98 % à 97 %.

3.4.4 Analyse par RCP d'urologie

Une analyse globale a été réalisée pour chacune des 9 RCP spécifiques d'urologie et d'une RCP polyvalente du département (Tableau 25). L'exhaustivité pour les cancers du rein était très variable d'une RCP à l'autre quelle que soit la source considérée. Par ailleurs pour une même RCP, l'exhaustivité variait également selon la source de données considérée. L'exhaustivité des trois sources variait de 97,3 % à 99,9 % selon la RCP.

I TABLEAU 25 I

Estimation de l'exhaustivité de chaque source pour les cancers du rein en fonction des RCP d'urologie ou polyvalente, Bouches-du-Rhône, 2013-2014

RCP Urologie	Revela13 (%)	PMSI (%)	ACP (%)	3 sources (%)
RCP Urologie - D	88,0	93,2	81,0	99,3
RCP Urologie - E	95,7	77,7	82,2	98,7
RCP Urologie - F	84,4	84,4	83,1	98,9
RCP Urologie - G	92,9	92,9	88,4	97,3
RCP Urologie - H	86,0	91,6	89,7	99,1
RCP Urologie - I	90,9	90,9	90,9	97,9
RCP Polyvalente - J	92,5	92,5	80,9	98,3
RCP Urologie - K	67,2	92,1	85,7	99,7
RCP Urologie - L	93,6	93,6	90,5	99,7
RCP Urologie - M	91,6	88,6	91,6	99,9

Lorsque l'on comparait l'exhaustivité pour les cancers du rein et celle pour les tumeurs urothéliales au sein d'une même RCP, l'exhaustivité pour les cancers du rein était généralement plus élevée pour la source Revela13, sauf pour les RCP Urologie D, F, K et L, contrairement à ce qui était observé pour les deux autres sources (meilleure exhaustivité du PMSI et des ACP pour les tumeurs urothéliales que pour les cancers du rein).

4. DISCUSSION

4.1. Principaux résultats

Dans cette étude, l'exhaustivité de l'observatoire Revela13 n'est pas ressortie homogène pour les trois cancers et variait globalement de 84 % pour les cancers des voies urinaires à 75 % pour les leucémies aiguës. Quel que soit le cancer étudié, les données du PMSI représentaient la source la plus exhaustive. L'exhaustivité des laboratoires ACP était soit identique, pour les tumeurs urothéliales, soit inférieure, pour le cancer du rein, à celle de l'observatoire Revela13.

Par ailleurs, quel que soit le cancer et la source de données considérés, les cas de cancers des patients âgés étaient moins bien identifiés que ceux des patients jeunes. C'est particulièrement marqué pour les leucémies aiguës pour lesquelles l'exhaustivité de la source Revela13 passait de 86 % chez les moins de 60 ans à 59 % pour les 80 ans et plus. La probabilité de capture pour cette classe d'âge était aussi diminuée dans la source PMSI, l'exhaustivité passant de 88 % à 67 %.

Pour les tumeurs urothéliales, le secteur de résidence et le stade de la tumeur étaient aussi des facteurs d'hétérogénéité. Les cas résidant hors de Marseille et ceux ayant une tumeur non infiltrante étaient moins bien identifiés. Pour ces tumeurs, les patients âgés résidant hors Marseille étaient moins bien identifiés, à la fois dans les sources Revela13 et DIM chez les hommes. Les femmes âgées de 85 ans et plus résidant hors Marseille étaient mal identifiées par la source Revela13.

Quel que soit le cancer, l'exhaustivité de Revela13 variait d'une RCP à l'autre, de 79 % à 87 % pour les leucémies aiguës, de 73 % à 95 % pour les tumeurs urothéliales et de 67 % à 96 % pour les cancers du rein. Après ajout des deux autres sources de données, l'exhaustivité estimée pour chaque RCP ne variait plus que de 93 % à 100% pour les trois cancers. L'apport des autres sources de données permettait ainsi de gommer les disparités géographiques mises en évidence tant au travers du secteur de résidence pour les tumeurs urothéliales qu'au travers de la RCP déclarante pour les trois cancers.

Avec les trois sources, l'exhaustivité globale de l'observatoire Revela13 variait de 95 % pour les leucémies aiguës à 99 % pour les tumeurs urothéliales, 97 % pour les cancers du rein.

4.2 Recueil des données

4.2.1 Extractions des fiches de RCP pour la complétude de l'observatoire

Le rattrapage de 15 à 20 % de cas supplémentaires grâce aux deux autres sources de données a montré que certains patients n'étaient pas identifiés dans les extractions de RCP reçues par l'observatoire Revela13, soit que les dossiers de ces patients n'avaient pas été présentés en RCP, soit que les fiches RCP n'avaient pas été transmises au réseau régional ou aux 3C ou saisies sur leur outil, soit encore que les requêtes d'extraction mises en œuvre n'avaient pas permis de les retrouver.

Les requêtes d'extraction mises en place auprès des différents partenaires ne concernaient que les nouveaux patients, ce qui n'a pas permis d'identifier les patients qui étaient passés plusieurs fois en RCP avant d'avoir un diagnostic de cancer confirmé. Ainsi, les tumeurs urothéliales ont pu ne pas être identifiées dans les extractions lorsqu'elles faisaient suite à plusieurs présentations en RCP de tumeurs antérieures superficielles (stade pTa). Cette situation étant considérée comme une récurrence, elle n'était pas retenue par les requêtes d'extraction. De même, les dossiers de patients présentés en RCP avec un cancer du rein

identifié radiologiquement mais non encore confirmé par histologie, n'étaient pas retenus dans ces requêtes, si le passage en RCP de la confirmation histologique mentionnait le statut de « rechute ».

Par ailleurs, le requêtage ciblé sur les RCP spécifiques d'urologie et d'hématologie n'a pas permis d'identifier des patients diagnostiqués dans les autres services et dont les dossiers étaient présentés dans d'autres RCP. Cette situation a été rencontrée notamment dans l'extraction du DIM de l'AP-HM où des cas de leucémies aiguës étaient hospitalisés en service de médecine interne.

Entre 2012 et 2015, plusieurs interruptions de transmissions des fiches des RCP au réseau régional OncoPaca-Corse ont été identifiées. Elles étaient dues à des problèmes de format informatique ou de modification du système informatique des établissements transmetteurs, mais aussi à des retards de saisies des fiches. En effet, lorsque les établissements ne saisissent pas directement leurs données en ligne dans l'outil OncoRCP, les problèmes de transfert de fichiers peuvent entraîner un retard dans l'intégration à la base régionale voir une perte de fiches RCP si celles-ci ne sont pas saisies rapidement.

4.2.2 Recueil des données des sources complémentaires

Le recueil des données complémentaires au travers des DIM et des laboratoires ACP a été compliqué et fastidieux malgré les autorisations Cnil et CCTIRS. Cette étude s'est déroulée sur plus de 2 ans et s'est heurtée aux mêmes difficultés que celles rencontrées par les registres, à savoir une réticence des laboratoires d'anatomopathologie ou des DIM à mettre leurs données à disposition [20]. Les raisons invoquées étaient le plus souvent l'absence de cadre juridique pour la transmission de données nominatives et les difficultés d'informer les patients, mais aussi l'absence de moyens alloués à ce travail supplémentaire.

Extraction des données des DIM

Afin de ne pas augmenter la charge de travail des DIM et de ne pas avoir à vérifier trop de cas non incidents, la sélection des cas a été réalisée uniquement sur le code du DP, sans prendre les diagnostics associés (DA). Ceci a probablement conduit à ne pas identifier certains cas, notamment les plus âgés présentant souvent de multiples comorbidités. Toutefois, certains DIM ont sélectionné leurs séjours à partir des DP et des DA et cela n'a pas permis de trouver des cas incidents supplémentaires.

L'application d'une période de « wash-out ^a » de 2 ans a été utile pour identifier des cas non incidents déjà hospitalisés pour les leucémies aiguës et les cancers du rein. En revanche, il n'a pas été possible d'identifier les tumeurs urothéliales dont le codage ne permettait pas pour la plupart des DIM, de différencier les séjours pour des tumeurs superficielles ou des lésions in situ, des séjours pour cancers invasifs. Certains DIM ont adopté une règle de codage des tumeurs non invasives en D41* versus des tumeurs invasives en C67 mais la plupart des établissements ont codé toutes les tumeurs urothéliales en C67 quel que soit leur degré d'invasion.

Les cancers du rein incidents ont été fréquemment identifiés par la succession de deux séjours séparés de 1 à 3 mois (le premier correspondant à la biopsie diagnostique et le second à l'intervention chirurgicale) mais le codage était variable d'une structure à l'autre. Certains établissements codaient le séjour correspondant à la biopsie avec le code D410 (tumeur d'évolution imprévisible du rein) et le séjour correspondant à l'intervention chirurgicale avec le code C64 alors que la plupart des établissements utilisaient le code C64 pour tous les séjours.

^a Nettoyage des données.

Les leucémies aiguës étaient fréquemment identifiées par le code générique de leucémie aiguë (C910, C920, C950), la mise à jour de la CIM10 concernant la différenciation de certaines leucémies était peu utilisée.

Extraction des données des ACP

La codification des diagnostics anatomopathologiques est recommandée par la spécialité mais n'est pas obligatoire. De plus, plusieurs nomenclatures de codage sont disponibles : Adicap la plus utilisée, CIMO3, CIM10 et Snowmed (thésaurus américain).

Dans cette étude, la plupart des laboratoires sollicités utilisaient la classification Adicap (8/13), deux laboratoires utilisaient la CIMO3 et/ou la CIM10, deux la classification Snowmed et un n'utilisait globalement pas de codage. De plus, au sein d'un même laboratoire, lorsqu'une classification était retenue, elle n'était pas systématiquement appliquée par l'ensemble des médecins du laboratoire. La diversité de nomenclature n'a pas permis la mise en œuvre d'une requête standardisée ce qui aurait facilité le recueil.

Deux laboratoires n'ont pas réussi à mettre en œuvre de requête informatique pour extraire les diagnostics pouvant répondre aux définitions de cas retenues. Pour ces laboratoires, le point d'entrée de la recherche de diagnostics n'était pas le patient mais le médecin prescripteur ce qui a sans doute généré la perte des diagnostics réalisés en 1^{ère} intention par d'autres médecins (oncologue, radiologues, etc.), notamment pour les biopsies de cancer du rein sous scanner.

Finalement, l'exploitation des données des laboratoires était donc très chronophage avec un nombre de cas retrouvé bien inférieur à celui des données du PMSI et un nombre de cas identifié uniquement par cette source qui ne représentait que 1,5 % de la totalité des cancers urinaires finalement inclus sur la période.

4.3 Méthode capture-recapture

Les conditions d'application de la méthode capture-recapture sont nombreuses et sont importantes à vérifier pour limiter les biais dans les résultats. La majorité de ces conditions étaient respectées, d'autres étaient contrôlées dans les analyses.

4.3.1 Vérification des hypothèses

L'identification de vrais cas dans chacune des 3 sources a reposé sur le retour au dossier systématique pour la vérification des cas identifiés.

L'identification des cas communs a fait l'objet de vérifications et d'exclusions successives. L'identification des cas communs a été réalisée à partir d'un algorithme croisant les identifiants des cas construits de la même manière pour chaque source. Cette détection a été complétée par une vérification des cas « non communs » pour retrouver d'éventuelles erreurs de saisies dans les variables identifiantes. Enfin, lors de la saisie de nouveaux cas dans la base de données, un algorithme de détection des doublons a permis de retrouver les cas qui auraient pu échapper aux vérifications précédentes. Disposer des mêmes variables identifiantes dans les trois sources n'est pas fréquent dans les études capture-recapture, mais facilite l'identification des cas communs, sans avoir à recourir à différents niveaux d'analyse avec des cas communs certains et d'autres probables comme dans l'étude sur les nouveaux diagnostics de VIH [19].

La période d'étude et la zone géographique étaient identiques pour les trois sources selon le protocole établi. Le code de résidence était toujours renseigné dans les données du PMSI

mais il n'était pas toujours présent dans les données des laboratoires ou des RCP. Pour ces deux sources les fiches sans code postal ont été recensées et vérifiées. Une erreur sur le code postal est toujours possible mais était limité par le géocodage à l'adresse, une erreur simultanée dans le libellé de l'adresse et le code postal était peu probable.

La nécessité d'une population close suppose qu'il n'y ait pas de mouvement de population (ni entrée, ni sortie) afin d'assurer une équiprobabilité de capture des cas dans chaque source. Dans cette étude portant sur des cas adultes, le diagnostic par ACP, le passage en RCP et le codage du séjour dans le PMSI ont été réalisés en général dans les 4 mois après le diagnostic ce qui a peu impacté la probabilité de capture. Seule la survenue d'un décès précoce a pu entraîner une non-homogénéité de capture par les différentes sources, avec un impact essentiellement pour les classes d'âge élevées .

4.3.2 Indépendance des sources

Comme le souligne Bailly [20], le diagnostic de cancer repose sur le CR anatomopathologique et la question de l'indépendance entre les sources peut se poser puisqu'au final la principale information est commune. Toutefois, les circuits d'information et l'organisation intrinsèque de chaque source sont a priori différents.

L'interaction retenue dans le modèle capture-recapture correspond souvent à la dépendance la plus forte [8]. Ces dépendances positives, bien que contrôlées dans les modèles, ont pu sous-estimer le nombre total de cas et donc augmenter l'exhaustivité estimée.

L'analyse capture-recapture est fréquemment utilisée pour étudier l'exhaustivité des registres mais présente des limites notamment celle de la dépendance entre sources qui est souvent présente en épidémiologie [21]. Dans notre étude, la dépendance positive a été effectivement vérifiée pour les tumeurs urothéliales pour chacune des sources prises deux à deux mais elle était du même niveau, expliquant ainsi que le modèle sans interactions a finalement été retenu. A l'inverse, pour les cancers du rein, le modèle retenu correspondait bien à la principale interaction mise en évidence entre les sources PMSI et ACP. En effet le diagnostic ACP est réalisé sur une pièce de néphrectomie d'un patient qui est bien évidemment hospitalisé, les cas de la liste ACP ont donc une forte probabilité d'être sur la liste PMSI. Il vaut donc mieux en épidémiologie faire des études de capture-recapture à 3 sources ou plus pour contrôler la dépendance et donner une estimation la moins biaisée possible du nombre de cas total.

4.3.3 Hétérogénéité de capture

Des variables d'hétérogénéité de capture avaient été identifiées *a priori*. Les analyses stratifiées menées sur ces variables ont mis en évidence des probabilités de capture de patients différentes selon ces variables.

Ainsi, quel que soit le cancer considéré, la probabilité de capture des patients était différente en fonction de l'âge. Pour les tumeurs urothéliales, d'autres variables telles que le secteur de résidence, le type et le degré de sévérité de la tumeur influent sur la probabilité de capture de la source Revela13 mais aussi, pour certaines d'entre elles, pour les deux autres sources. Ces hétérogénéités de capture ont pu remettre en cause les résultats de l'analyse globale, pour chacun des 3 cancers. Cependant, la somme des estimations par strates était très proche de l'estimation obtenue par le modèle global, quel que soit le cancer étudié, ce qui minimise le risque que ces hétérogénéités soient source de biais dans l'estimation globale.

4.4 Estimations de l'exhaustivité de l'observatoire Revela13

Compte tenu de l'origine des signalements de Revela13, soit au travers des extractions de RCP soit à partir des dossiers présentés en RCP, l'exhaustivité de Revela13 reflète celle des RCP et les résultats observés peuvent donc être comparés à ceux des études d'exhaustivité menées depuis leur mise en place.

4.4.1 Exhaustivité des cancers des voies urinaires

Une exhaustivité supérieure à 80 %

L'exhaustivité de l'observatoire Revela13 estimée pour les cancers urologiques, hors prostate, était supérieure à celle observée dans le rapport de la Haute autorité de santé (HAS) de 2016 [23], pour l'uro-néphrologie qui était à 78 %. Dans cette étude, l'uro-néphrologie était une des trois disciplines avec le moins bon taux d'exhaustivité des RCP. En revanche, l'exhaustivité de l'observatoire Revela13 était similaire à celle estimée par le réseau OncoBretagne en 2015 (82,8 %) pour les RCP d'urologie [24].

Une influence des territoires et des pratiques des praticiens

L'étude d'OncoBretagne a montré que les taux d'exhaustivité de passage en RCP étaient dépendant des territoires étudiés et que les dossiers qui ne passaient pas en RCP étaient le plus souvent des cancers de vessie de bas grade. Dans notre étude, la dépendance en fonction du territoire était également vérifiée pour les tumeurs urothéliales. La probabilité de capture par Revela13 était en effet plus élevée pour les patients résidant à Marseille par rapport aux autres patients. Ce résultat était par ailleurs également observé pour les sources DIM et ACP.

Un autre résultat confirme l'influence du territoire sur la probabilité de capture. En effet, parallèlement à l'étude globale d'exhaustivité, l'étude par RCP pour les 10 RCP d'urologie ou polyvalente du département des Bouches-du-Rhône montrent des différences importantes entre RCP. Ainsi, l'exhaustivité de l'observatoire Revela13 estimée pour chaque RCP spécifique variait de 72 à 96 % pour les cancers du rein et de 77 à 93 % pour les tumeurs urothéliales. Ces résultats reflètent vraisemblablement des différences dans l'organisation des RCP mais s'expliquent aussi par les interruptions de transmission des fiches de RCP. Par ailleurs, il a été parfois difficile de reconstituer précisément les cohortes de patients discutés dans une RCP. En effet, les dossiers de certains patients avaient pu être présentés dans plusieurs RCP, d'une part car les patients pouvaient avoir consulté sur plusieurs sites (2^{ème} avis de spécialistes) et d'autre part, parce que de nombreux spécialistes privés consultent dans plusieurs établissements et peuvent être amenés à participer à deux RCP spécifiques. L'exercice en multisite est plus fréquent dans les zones en dehors de Marseille. Dans les petites cliniques privées, situées principalement en dehors de Marseille, les urologues ne disposent pas d'un secrétariat spécifique qui collige les dossiers de patients à présenter en RCP et peuvent donc parfois omettre de présenter un dossier. Pour autant, si ces hypothèses peuvent expliquer la moindre probabilité de capture par l'observatoire Revela13 des patients résidant en dehors de Marseille, il est difficile d'interpréter ces différences pour les données du PMSI et des ACP retrouvées pour les tumeurs urothéliales mais pas pour les cancers du rein.

Une influence de l'âge confirmée

L'influence de l'âge sur la probabilité de capture de l'Observatoire était attendue. En effet, elle avait déjà été mise en évidence dans une étude d'exhaustivité et de qualité des RCP pour les cancers prostatiques [25]. Dans cette étude, la probabilité de passage en RCP diminuait

fortement chez les cas âgés de 85 ans et plus comparativement à ceux de la classe d'âge 70–74 ans. Avec l'âge, le risque de développer une tumeur augmente. Par ailleurs, notamment pour les tumeurs urothéliales, la probabilité que des polypes superficiels déjà connus dégénèrent en stade invasif augmente également chez le sujet âgé. Ainsi, plus le patient est âgé, plus la probabilité que son dossier ait déjà été discuté en RCP pour des lésions superficielles est élevée. Il est possible que certains patients, dits en récurrence, puissent échapper à une nouvelle présentation en RCP.

Cette raison pourrait également expliquer la moindre probabilité de capture par l'observatoire des TVNIM qui font souvent suite à une ou plusieurs récurrences rapprochées de tumeurs superficielles. De plus, la prise en charge sur le plan thérapeutique des TVNIM est similaire à celle des tumeurs superficielles ne nécessitant pas un passage en RCP pour discuter de l'orientation thérapeutique. Comme le rappelle Guillem *et al.*, la RCP a pour vocation première d'améliorer la prise en charge du patient en discutant les différentes orientations médicales proposées [26].

Une exhaustivité plus faible pour les cancers du rein en ACP et PMSI

Lorsque l'on s'intéresse à l'exhaustivité des cancers du rein, on note que, si elle est globalement identique à celle des tumeurs urothéliales pour Revela13, elle est plus faible pour les deux autres sources. Ainsi, pour la source ACP, l'exhaustivité des tumeurs du rein est inférieure à 80 % alors qu'elle est de 86 % pour les tumeurs urothéliales et respectivement de 86 % et 92 % pour la source PMSI. La sous-estimation possible des cas de cancers du rein est sans doute liée aux extractions ciblées sur les urologues partenaires qui ne permettent pas d'identifier les CR de biopsie sous scanner réalisés par des radiologues qui n'adhèrent pas au réseau de l'observatoire Revela13 et qui de fait, ne lui ont pas transmis les CR. Les cas de cancers du rein non opérés par la suite (surveillance IRM, traitement par radiofréquence ou cryothérapie) ne peuvent pas être identifiés par la source ACP. La forte dépendance existante entre les sources ACP et PMSI pour les cancers du rein fait que les cas non détectés par les ACP diminuent la probabilité d'être identifiés par les données du PMSI.

4.4.2 Exhaustivité des leucémies aiguës

Une exhaustivité faible

L'exhaustivité de l'observatoire pour les leucémies aiguës est faible avec 75 %, bien en deçà de ce qui a pu être observé dans l'étude de la HAS (85 %).

L'exhaustivité de l'observatoire Revela13 est très liée à celle de la RCP d'hématologie de l'IPC, puisque 60% des cas incidents de leucémies aiguës sont diagnostiqués à l'IPC. Or, lors des analyses réalisées pour chacune des RCP spécifiques, il s'est avéré que même si celle-ci était supérieure à l'exhaustivité globale, elle n'était que de 82 %. De plus, l'exhaustivité de cette RCP spécifique a diminué entre 2013 et 2014 alors que le taux de signalement direct par les hématologues est passé de 43 % en 2013 à 12 % en 2014 illustrant l'importance de ces signalements directs pour une meilleure détection des leucémies aiguës (résultats non présentés). Une base de données spécifique concernant les leucémies aiguës est renseignée par le service d'hématologie de l'IPC. Cette base ne contient pas toutes les variables identifiantes nécessaires mais elle va être consultée systématiquement pour compléter le recueil sur cette structure.

Une influence de l'âge importante

Avec l'âge le risque de développer une leucémie aiguë augmente. Une rapide analyse des dossiers de LAM n'ayant pas de fiches RCP et détectés par la source PMSI a montré que les cas très âgés transférés en soins palliatifs ou décédés rapidement au décours de

l'hospitalisation n'étaient pas discutés en RCP. Les cas décédés lors de la chimiothérapie d'induction mise en œuvre au cours du séjour de découverte de la maladie n'avaient également pas systématiquement de fiche RCP dans le dossier informatique et ce, indépendamment de l'âge.

5. CONCLUSIONS ET PERSPECTIVES

Cette étude a permis d'évaluer l'exhaustivité de l'observatoire Revela13, mais aussi des sources complémentaires. L'ajout des données du PMSI et, dans une moindre mesure de celles des ACP, permettent d'améliorer nettement l'exhaustivité de l'observatoire qui augmente de 75 % à 95 % pour les leucémies aiguës, de 84 % à 99 % pour les tumeurs urothéliales et de 84 % à 97 % pour les cancers du rein.

Malgré l'incitation à présenter tous les nouveaux cas de cancer en RCP, l'exhaustivité de la source Revela13 seule n'est pas suffisante car elle ne permet pas de s'assurer de l'homogénéité de cette exhaustivité sur le territoire départemental. Pour autant, le mode de fonctionnement de l'observatoire est intéressant car il permet à ce dernier d'être bien accepté par les médecins spécialistes. Le recours aux fichiers de RCP comme source principale de données permet un travail en termes de coût/efficacité bien meilleur que pour les deux autres sources de données. Ces fichiers permettent de faire un premier tri entre tumeurs cancéreuses et non cancéreuses mais aussi, même si ce critère reste à améliorer, entre cas incident et cas prévalent. De plus, la validation des cas par retour au dossier est facilitée et raccourcie par la sollicitation directe des réseaux de spécialistes. Notre étude indique aussi que la performance en termes d'exhaustivité du PMSI seul est meilleure que celle de l'observatoire Revela13, résultat d'autant plus intéressant que la base de données est disponible directement. Cependant, la distinction entre cas incident et prévalent ainsi qu'entre tumeur urothéliale invasive et non invasive, reste difficile.

Toutefois, dans une perspective de poursuite de l'observatoire, la source PMSI, compte-tenu de sa contribution majoritaire à l'amélioration de l'exhaustivité, pourrait être ajoutée en routine au fonctionnement de l'observatoire Revela13, sans que cela n'engendre une charge de travail trop importante. En effet, les extractions des données du PMSI ont l'avantage d'être standardisées et facilement intégrables dans la base Revela13 pour un traitement automatisé des données. De plus, les résultats d'exhaustivité montrent que les variations entre les sources Revela13 et PMSI pour chaque RCP ne vont pas toujours dans le même sens, illustrant la complémentarité de l'association de ces deux sources de données pour gommer l'hétérogénéité de recueil sur le territoire. La contribution des données des laboratoires ACP est moindre et l'absence de standardisation au niveau de leur production rend impossible leur utilisation en routine pour l'instant. Ces données peuvent toutefois constituer une source ponctuelle pour retrouver des cas diagnostiqués par des spécialistes travaillant sur plusieurs sites.

La correction de l'hétérogénéité de recueil des données issues de Revela13, à partir des deux sources complémentaires, va permettre la mise en œuvre des analyses spatiales et de la recherche de clusters à l'échelon de la commune pour le département des Bouches-du-Rhône.

5.1 Amélioration de l'exhaustivité de l'observatoire Revela13

5.1.1 Optimiser les requêtes d'extraction des RCP

Plusieurs points d'améliorations ont été identifiés par l'étude d'exhaustivité. En premier lieu, la prise en compte de l'ensemble des dossiers discutés en RCP, et non plus uniquement ceux correspondant à un premier passage en RCP ou à un nouveau patient, devrait permettre d'améliorer l'identification des cas incidents et notamment ceux qui passent en RCP avant d'avoir un diagnostic de cancer confirmé par histologie. Ces cas seront ensuite suivis jusqu'à la confirmation histologique nécessaire à leur inclusion.

L'extraction doit porter sur toutes les RCP et pas seulement sur les RCP d'urologie et d'hématologie. Des dossiers de cas de cancer du rein hospitalisés à l'origine pour exploration de douleurs abdominales en service de gastroentérologie ont été présentés en RCP de pathologies digestives, par exemple.

Les requêtes d'extraction des données de RCP sont modifiées régulièrement pour prendre en compte les évolutions concernant le codage des maladies, les pratiques des cliniciens en RCP et les modifications liées à la gestion des bases de données. Ces modifications sont parfois trop tardives, nécessitant une extraction supplémentaire rétrospective sur plusieurs années avec un recueil des données plus compliqué et un nombre de dossiers non retrouvés plus important. Un échange plus régulier et formalisé avec les responsables du réseau régional et des 3C concernés serait utile pour prendre en compte rapidement les différentes modifications mises en œuvre. De même, la transmission, en même temps que le fichier trimestriel de données, d'un historique des interruptions de transmission permettrait à l'équipe projet d'adapter ses modalités de recueil pour les RCP concernées de façon à éviter la perte de cas incidents.

La nouvelle requête du réseau OncoPaca-Corse tenant compte des éléments précédents a été testée de manière rétroactive sur la période 2012-2014. Sur les 234 cas de cancers des voies urinaires retrouvés par l'étude d'exhaustivité, 72 cas (30 %) ont été identifiés par la nouvelle extraction. En revanche, aucun des douze cas de leucémies aiguës identifiés par le PMSI n'a été retrouvé dans cette extraction.

La nouvelle requête élaborée avec le 3C de l'AP-HM prend en compte l'ensemble des dossiers discutés dans les 3 RCP sélectionnées sans ajout de nouvelles RCP. Elle a été testée sur les cas diagnostiqués entre 2015 et 2016 dans les deux RCP d'urologie de l'AP-HM, mais n'a pas encore été testée sur la RCP d'hématologie et elle a d'ores et déjà montré une meilleure détection des cas de cancers des voies urinaires.

La nouvelle requête du 3C de l'IPC tenant compte des éléments précédents a été testée de manière rétroactive sur la période allant de mai 2013 à décembre 2014. Les données antérieures à avril 2013 n'ont pas pu être vérifiées suite au changement, à cette date, du système d'exploitation interne à l'établissement. Sur les 12 patients atteints d'un cancer des voies urinaires identifiés par l'étude d'exhaustivité, avec un diagnostic postérieur au 1^{er} mai 2013, 10 ont été retrouvés dans la nouvelle extraction. Concernant les leucémies aiguës, parmi les 18 patients identifiés par le DIM de l'IPC, avec un diagnostic postérieur au 1^{er} mai 2013, 5 patients ont été retrouvés dans l'extraction des RCP, 12 patients n'avaient pas de fiches RCP dans leurs dossiers et 1 patient avait bien une fiche RCP qui n'avait pas pu être sélectionnée par la nouvelle extraction. Ces résultats illustrent d'ores et déjà que les nouveaux critères de requête améliorent la sélection des cas de cancers des voies urinaires mais qu'un travail reste à faire pour mieux identifier les leucémies aiguës.

5.1.2 Compléter le recueil des données de RCP

Si l'optimisation des requêtes d'extraction des RCP permet d'améliorer l'exhaustivité de l'observatoire Revela13, elle n'est pour autant pas suffisante pour assurer seule l'homogénéité de recueil sur l'ensemble du département et devra donc être complétée par d'autres données.

Les données du PMSI

Dans cette étude, la source PMSI est la plus exhaustive quel que soit le cancer étudié. La prise en compte systématique des données du PMSI permettra de capter une grande part des cas non identifiés par l'observatoire Revela13. Après l'obtention d'une modification de l'accord Cnil et de l'accord des directeurs d'établissements et des responsables de DIM, la mise en

œuvre de ce recueil pour les données enregistrées à partir de 2017 pourrait être relativement simple.

Les données extraites par une requête d'extraction standardisée peuvent être obtenues sous format informatique facilement téléchargeable dans la base de données Revela13. Les données pourront être déposées via la plateforme sécurisée Epifiles dans un entrepôt dédié à chaque DIM. Le traitement des données sera similaire à celui mis en œuvre pour les données issues des RCP. Une concertation sera organisée avec les différents DIM pour optimiser la requête d'extraction en fonction de leur expérience et pour orienter vers une transmission en routine.

Augmenter les signalements en développant le réseau des praticiens partenaires

Intégrer les oncologues et les radiothérapeutes au travers de leur participation aux RCP spécifiques, permettra d'identifier des cas supplémentaires, notamment quand les urologues ayant une activité multisite ne peuvent pas participer régulièrement à toutes les RCP les concernant.

Pour les tumeurs du rein, il serait intéressant d'associer les radiologues puisque ce sont eux qui réalisent les biopsies dirigées sous examen scanographique. Bien qu'en général les résultats anatomopathologiques sont transmis au médecin prescripteur, nous avons pu identifier, par ce biais, des patients suivis par IRM et pas toujours présentés en RCP ou des patients non traités ultérieurement par un des chirurgiens de notre réseau d'urologue (traités probablement hors du département).

Ajouter une source de données biologiques pour les leucémies aiguës.

L'exhaustivité des leucémies aiguës est bien améliorée par l'ajout des données issues du PMSI mais il semble nécessaire de trouver une solution pour recueillir des données biologiques de confirmation des cas de leucémies aiguës.

5.1.3 Mieux cibler les cancers des personnes âgées

Sensibiliser les participants aux RCP

Quel que soit le cancer étudié, l'exhaustivité de l'observatoire Revela13 diminue nettement pour les cas les plus âgés. Plusieurs explications sont possibles comme la difficulté de mettre en œuvre un traitement chez des patients souvent polyopathologiques ou le transfert directement en soins palliatifs de certains patients diagnostiqués tardivement. Les praticiens du réseau devront être sensibilisés à ce problème lors de la restitution des résultats de cette étude par l'équipe projet aux différentes RCP. Des pistes d'améliorations pour la prise en compte des cas très âgés seront étudiées en lien avec les 3C et le réseau OncoPaca-Corse.

Ajouter les données de réunions pluridisciplinaires en oncogériatrie

Quelle que soit la pathologie étudiée, les personnes âgées de plus de 85 ans pour l'urologie ou de plus de 80 ans pour les leucémies aiguës sont moins souvent présentes dans la base du réseau OncoPaca-Corse et, de fait, moins souvent incluses dans l'observatoire Revela13. La moindre présentation des dossiers des sujets âgés en RCP est rapportée dans plusieurs études [19, 23]. Il est possible que l'organisation actuelle des RCP ne réponde pas aux besoins des praticiens pour cette population âgée dont la prise en charge est souvent guidée par le souci de ne pas sur-traiter un patient considéré comme à risque.

La mise en place des réunions de concertation en oncogériatrie (RCOG), plus adaptées à ce type de patients, car visant à identifier ceux qui devraient bénéficier d'une évaluation

gériatrique permettant la modulation du traitement (allègement ou renforcement du protocole), pourrait constituer une réponse à cette problématique [32].

L'unité de coordination en oncogériatrie (Ucog) PACA-Ouest organise une réunion bimensuelle de concertation en oncogériatrie impliquant plusieurs établissements des Bouches-du-Rhône et du Vaucluse. Les données des RCOG pourraient être recueillies en complément des données de RCP.

Prendre en compte des diagnostics associés dans l'extraction des données du PMSI.

Les patients âgés présentent fréquemment plusieurs comorbidités associées et même si un diagnostic de cancer est effectué au cours du séjour hospitalier, une autre pathologie également aggravée au cours du séjour peut être codée en diagnostic principal. La sélection des séjours sur le diagnostic principal et les diagnostics associés devraient permettre de mieux identifier ces cas.

5.2 Compléter l'étude d'exhaustivité sur les années 2015-2016

L'étude d'exhaustivité a mis en évidence des disparités de recueil de données au niveau des RCP partenaires qui se traduisent par une hétérogénéité spatiale de ce recueil à l'échelle du département. Cette étude a permis de corriger ces disparités. Par ailleurs, les premières analyses spatiales réalisées sur les données consolidées 2013-2014, ont montré que les données actuelles n'étaient pas suffisamment robustes pour conclure sur les résultats de ces analyses. Il est donc indispensable de disposer d'un nombre de données plus important en ajoutant les cas diagnostiqués sur la période 2015-2016 et en s'assurant de l'exhaustivité de recueil de ces données sur cette période.

Une demande d'extraction des données 2015-2016 a d'ores et déjà été adressée aux DIM partenaires pour compléter les données recueillies sur cette période par l'observatoire Revela13.

Les données du PMSI seront complétées par des extractions ciblées des données ACP auprès des laboratoires travaillant avec les urologues qui interviennent sur plusieurs sites et/ou avec les praticiens identifiés comme peu déclarants lors de la première étude d'exhaustivité.

Une nouvelle analyse capture-recapture sur les données 2013-2016 permettra d'évaluer l'évolution de l'exhaustivité de l'observatoire Revela13 au fur et à mesure de son implantation dans le département mais aussi d'évaluer les améliorations des requêtes d'extraction des bases RCP auprès des différents partenaires.

Références bibliographiques

1. Classification statistique internationale des maladies et des problèmes de santé connexes. Dixième révision. OMS. 1993.
2. Thésaurus de la codification Adicap. Edition 2009. Version 5.04. Disponible à l'URL : www.adicap.asso.fr
3. Gallay A, Nardone A, Vaillant V, Desenclos JC. La méthode capture-recapture appliquée à l'épidémiologie : principes, limites et applications. *Revue d'épidémiologie et de santé publique*, 2002;(50):219-32.
4. Berger F, Parent du Châtelet I, Bernillon P et Gallay A. Surveillance des infections invasives à méningocoque en France métropolitaine en 2005 – Évaluation quantitative par la méthode de capture-recapture à trois sources. Saint-Maurice (Fra) : Institut de veille sanitaire, août 2010, 43 p. Disponible à partir de l'URL : www.santepubliquefrance.fr
5. Loury P, Jones G, Chappert JL, Pivette M, Hubert B et le groupe Ecsir. Analyse de l'exhaustivité et de la qualité de la surveillance des gripes sévères, 2009-2013. Saint-Maurice : Santé publique France, 2017, 59 p. Disponible à partir de l'URL : www.santepubliquefrance.fr
6. Plouvier S, Bernillon P, Buisset E, Martin P, Benoit E, Leroy X, Gallay A, Ligier K. Évaluation de l'exhaustivité de la première année d'enregistrement du registre général des cancers de Lille et de sa région. Congrès SFSP, 2011, Lille. Disponible à l'URL : <http://www.registrecancers59.fr/>
7. Guillem P, Lavergne C, Cans C, Comment évaluer l'exhaustivité d'un registre de morbidité ? l'exemple du registre des handicaps de l'enfant et observatoire périnatal de l'Isère. *RESP* 200 ;48 :41-51.
8. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev* 1995;17(2):243-64.
9. Capture-recapture and multiple-record systems estimation I: history and theoretical development. International working group for disease monitoring and forecasting. *Am J Epidemiol* 1995;142(10):1047-58.
10. Capture-recapture and multiple-record systems estimation II: applications in human diseases. International working group for disease monitoring and forecasting. *Am J Epidemiol* 1995;142(10):1059-68.
11. Arber DA *et al.* The 2016 revision to the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia. *Blood* 2016; 127(20):2391-405.
12. Eble J. World Health Organisation classification of tumors. Pathology and genetics of tumors of the urinary system and male genital organs. Lyon: IARC Press, 2004.
13. Moch H, Cubilla A L, Humphrey PA, Reuter VE & Ulbright TM. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *Eur. Urol* 2016;70:93-105.
14. Lamy S, Daubisse L, Lagadic C, Godddard J, Tourniaire C, Delpierre E *et al.* Exhaustivité du passage en RCP : étude dans un département d'Occitanie disposant d'un registre des cancers, le Tarn. *RESP* 2017;65(supp 2):S82.
15. Raffoux E, Dombret H. Leucémies aiguës du sujet âgé. *Correspondances en Onco-hématologie* 2008 ;3(1) :36-42.

16. Wittes JT, Colton T, Sidel VW. Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *J Chronic Dis* 1974;27(1):25-36.
17. Bishop YMM, Fienberg SE, Holland PW. Estimating the size of a closed population. *Discrete Multivariate Analysis: Theory and Practice*. The MIT Press, 1975. p. 229-56.
18. R software : Package Rcapture. Disponible à partir de l'URL : <https://cran.r-project.org/web/packages/Rcapture/Rcapture.pdf>.
19. Estimation du nombre de diagnostics d'infection par le VIH chez l'enfant par la méthode capture-recapture, Francemétropolitaine, 2003-2006 – Rapport de master. Saint Maurice (Fra) : Institut de veille sanitaire, juillet 2010, 27p. Disponible à partir de l'URL : www.santepubliquefrance.fr
20. Bailly L. Validation et exploitation d'un registre histologique des cancers. Estimation par capture-recapture de l'exhaustivité par modèles log-linéaires et selon les modèles écologiques Mtbh bayésiens. *Biostatistique épidémiologie*. Université Montpellier I, 2011
21. Tilling K. Capture-recapture methods - useful or misleading?. *Int J Epidemiol* 2001;30:12-14.
22. Comité national des registres. Registres épidémiologiques et accès aux sources de données standardisées : état des lieux et perspectives d'amélioration. Consulté sur internet le 02/11/2011 : http://www.sante.gouv.fr/IMG/pdf/Contribution_du_Comite_national_des_registres_SpFrance-Inserm_-_propositions.pdf
23. Réunion de concertation pluridisciplinaire en cancérologie – Résultats nationaux de la campagne 2016 – Données 2015. HAS Décembre 2016 34p. Disponible à partir de l'URL : https://www.has-sante.fr/portail/upload/docs/application/pdf/2016-12/rapport_2016_rcp_vf.pdf
24. Exhaustivité des passages en RCP d'urologie Onco Bretagne. Décembre 2015. Disponible à partir de l'URL : https://www.oncobretagne.fr/wp-content/uploads/2016/01/Synth%C3%A8se_enqu%C3%AAtes_urologie_2015.pdf
25. Huo Yung Kai S, Delpierre C, Gaudin C, Goddard J, Daubisse-Marliac L, Soulié M *et al*. Exhaustivité et qualité des réunions de concertation pluridisciplinaire ; l'exemple des cancers prostatiques en Midi-Pyrénées. *Progrès en urologie* 2011;21:879-86.
26. Guillem P, Bolla M, Courby S, Descotes JL, Laramas M, Moro-Sibilot D. Évaluation des réunions de concertation pluridisciplinaire en cancérologie : quelles priorités pour quelles améliorations ? *Bulletin du Cancer* 2011;98(9):989-998

ANNEXES

Annexe 1 : Fiches de recueil

Annexe 2 : Libellés des codes CIM10

Annexe 3 : Liste des établissements hospitaliers

Annexe 4 : Liste des laboratoires d'anatomopathologies et laboratoires d'hématologies sollicités pour l'étude d'exhaustivité

Annexe 5 : Liste des codes Adicap

Annexe 6 : Schéma de croisement des données prévu dans le protocole

Annexe 7 : Rappel sur la méthode capture-recapture

Annexe 8 : R-Capture

Annexe 9 : Analyse de sensibilité avec l'estimateur de Chao

Annexe 10 : Analyse stratifiée des tumeurs urothéliales

Annexe 1 : Fiches de recueil

	<h3>FICHE DE SIGNALEMENT</h3> <h3>CANCER UROTHÉLIAL PRIMAIRE</h3> <p>Définition de cas : patient âgé de plus de 18 ans, résidant dans les Bouches-du-Rhône, avec un 1^{er} diagnostic de tumeur invasive de la vessie (pT1), de l'uretère ou du bassinot</p>
---	---

A compléter en ligne sur : <https://voozanoo.invs.sante.fr/revela13>

Ou à renvoyer complétée par courrier/ fax ou mail avec la mention « secret médical » à :

Dr Laurence PASCAL – Observatoire REVELA 13- ARS PACA – CIRE Sud/Bureau 526 - 132 Boulevard de Paris - CS 50039-13331 Marseille cedex 3 - Tél. 04.13.55.83.11 - Fax. 04.13.55.83.47-mèl : laurence.pascal@ars.sante.fr

Informations patient	
Nom :	Prénom :
Nom de jeune fille :	Sexe : Masculin <input type="checkbox"/> / Féminin <input type="checkbox"/>
Date de naissance : ___ / ___ / _____	Commune de naissance :
Adresse :	
Code postal :	Commune : Tél : / / / / / / /
Décédé : oui <input type="checkbox"/> non <input type="checkbox"/> ne sait pas <input type="checkbox"/>	Date de décès : ___ / ___ / _____
Numéro identifiant (à compléter par l'équipe projet) :	

volet détachable

Signalement	
Date de signalement : ___ / ___ / 20___ <input type="checkbox"/> Direct <input type="checkbox"/> Ext. ONCOPACA <input type="checkbox"/> 3C <input type="checkbox"/> Autre	
Médecin déclarant	Nom : Hospitalier <input type="checkbox"/> Libéral <input type="checkbox"/>
	Établissement ou commune d'exercice : Tél : / / / / / / /
Médecin traitant	Nom : Tél : / / / / / / /
	Adresse :
Autre médecin référent	Nom : Tél : / / / / / / /
	Établissement ou commune d'exercice :

Renseignements cliniques	
Date de début de la maladie : ___ / ___ / 20___	
Siège de la tumeur (code CIM-10) : C / ___ / ___ / ou Libellé en clair :	
Base diagnostic : <input type="checkbox"/> cytoscopie <input type="checkbox"/> cytologie <input type="checkbox"/> histologie <input type="checkbox"/> certificat décès Date diagnostic : ___ / ___ / 20___	

Données anatomopathologiques	
Date intervention : ___ / ___ / 20___	
Type de la tumeur : Unifocale <input type="checkbox"/> Multifocale <input type="checkbox"/>	
Stade pTNM :	Grade :
Taille de la tumeur principale (en mm) :	
Numéro du compte-rendu anatomopathologique :	
Cabinet ayant réalisé l'examen anatomopathologique :	

Numéro identifiant (à compléter par l'équipe projet) :



FICHE DE SIGNALEMENT CANCER INCIDENT DU REIN

Définition de cas : patient âgé de plus de 18 ans, résidant dans les Bouches-du-Rhône, avec un 1^{er} diagnostic de tumeur invasive du rein (hors voie excrétrice supérieure)

A compléter en ligne sur : <https://voozanoo.invs.sante.fr/revela13>

Ou à renvoyer complétée par courrier/ fax ou mail avec la mention « secret médical » au Dr Laurence PASCAL : Observatoire REVELA 13- ARS PACA – CIRE Sud/Bureau 526 - 132 Boulevard de Paris - CS 50039 - 13331 Marseille cedex 3 Tél. 04.13.55.83.11 - Fax. 04.13.55.83.47-mél : laurence.pascal@ars.sante.fr

Informations patient

Nom :	Prénom :	
Nom de jeune fille :	Sexe : Masculin <input type="checkbox"/> / Féminin <input type="checkbox"/>	
Date de naissance : / /	Commune de naissance :	
Adresse :		
Code postal :	Commune :	Tél : / / / / / /
Décédé : oui <input type="checkbox"/> non <input type="checkbox"/> ne sait pas <input type="checkbox"/>	Date de décès : / /	
Numéro identifiant (à compléter par l'équipe projet) :		

volet détachable

Signalement

Date de signalement : ____ / ____ / 20____		<input type="checkbox"/> Direct	<input type="checkbox"/> Ext. ONCOPACA	<input type="checkbox"/> 3C	<input type="checkbox"/> Autre
Médecin déclarant	Nom :	Hospitalier <input type="checkbox"/> Libéral <input type="checkbox"/>			
	Établissement ou commune d'exercice :				
Médecin traitant	Tél : / / / / / /				
	Nom :	Tél : / / / / / /			
Autre médecin référent	Adresse :				
	Nom :	Tél : / / / / / /			
Établissement ou commune d'exercice :					

Renseignements cliniques

Date de début de la maladie : ____ / ____ / 20____	
Siège de la tumeur (code CIM-10) : C/____ / ____ / ____ ou Libellé en clair :	
Base diagnostic : <input type="checkbox"/> Imagerie <input type="checkbox"/> histologie <input type="checkbox"/> certificat décès	Date imagerie : ____ / ____ / 20____

Données anatomopathologiques

Date de la Biopsie : ____ / ____ / 20____	Date de l'intervention : ____ / ____ / 20____
Histologie : <input type="checkbox"/> cellules conventionnelles / <input type="checkbox"/> tubulo-papillaires / <input type="checkbox"/> chromophobes / <input type="checkbox"/> sarcomatoïdes / <input type="checkbox"/> autre	
Stade pTNM :	Grade de Fuhrman :
Taille de la tumeur principale (en mm) :	
Numéro du compte-rendu anatomopathologique :	
Cabinet ayant réalisé l'examen anatomopathologique :	
.....	

Numéro identifiant (à compléter par l'équipe projet) :



FICHE DE SIGNALEMENT LEUCEMIE AIGUE DE L'ADULTE

Définition de cas : patient âgé de plus de 18 ans, résidant dans les Bouches-du-Rhône, avec un 1^{er} diagnostic de leucémie aiguë ou de lymphome inclus dans la classification OMS 2008

A compléter en ligne sur : <https://voozanoo.invs.sante.fr/revela13>

Ou à renvoyer par courrier/ fax ou mail avec la mention « secret médical » au Dr Laurence PASCAL :
ARS PACA – CIRE Sud/Bureau 526 - 132 Boulevard de Paris - CS 50039 - 13331 Marseille cedex 3
Tél. 04.13.55.83.11 - Fax. 04.13.55.83.47- mèl : laurence.pascal@ars.sante.fr

Informations patient

Nom : Prénom :
Nom de jeune fille : Sexe : Masculin / Féminin
Date de naissance : ___ / ___ / _____ Commune de naissance :
Adresse :
Code postal : Commune : Tél : ___ / ___ / ___ / ___ / ___ / ___
Décédé : oui non ne sait pas Date de décès : ___ / ___ / _____
Numéro identifiant (à compléter par l'équipe projet) : _____
volet détachable

Signalement

Date de signalement : ___ / ___ / _____ Direct Ext. ONCOPACA 3C Autre

Médecin déclarant	Nom : Hospitalier <input type="checkbox"/> Libéral <input type="checkbox"/>
	Établissement ou commune d'exercice : Tél : ___ / ___ / ___ / ___ / ___ / ___
Médecin traitant	Nom : Tél : ___ / ___ / ___ / ___ / ___ / ___
	Adresse :

Renseignements cliniques

Date de début de la maladie : ___ / ___ / 20___
Libellée classification OMS 2008 : _____
Code CIM-10 : C / ___ / ___ / ___ / ___ / ___ CIMO3 : _____
Antécédent de SMD Oui Non Ne sait pas Date Diagnostic SMD : ___ / ___ / 20___
Antécédent LLC Oui Non Ne sait pas Antécédent LMC Oui Non Ne sait pas
Antécédents de cancer Oui Non Ne sait pas
Antécédent de chimio anticancéreuse Oui Non Ne sait pas
Antécédents d'irradiations Oui Non Ne sait pas

Données anatomopathologiques

Date du diagnostic cytologique : ___ / ___ / 20___
Type LAM LAL LA Non classable Lymphome Burkitt Lymphome lymphoblastique
Laboratoire ayant réalisé l'examen cytologique :
Cytogénétique Oui Non NSP

Histoire professionnelle

Actif Retraité Sans emploi Ne sait pas
Dernier emploi ou emploi actuellement occupé :
Profession : Secteur d'activité :
Entreprise :
Emploi le plus longuement occupé dans la carrière :
Profession : Secteur d'activité :
Entreprise :

Numéro identifiant (à compléter par l'équipe projet) : _____

Annexe 2 : Libellés des codes CIM10

Codes CIM10 tumeurs urothéliales et cancers du rein

- C64 – Tumeur maligne du rein, à l'exception du bassinet
- C65 – Tumeur maligne du bassinet
- C66 – Tumeur maligne de l'uretère
- C67 – Tumeur maligne de la vessie
- C68 – Tumeur maligne des organes urinaires autres et non précisés
- D41 - Tumeur des organes urinaires à évolution imprévisible ou inconnue
- D410 - Tumeur à évolution imprévisible ou inconnue du rein
- D411 - Tumeur à évolution imprévisible ou inconnue du bassinet
- D412 - Tumeur à évolution imprévisible ou inconnue de l'uretère
- D413 - Tumeur à évolution imprévisible ou inconnue de l'urètre
- D414 - Tumeur à évolution imprévisible ou inconnue de la vessie
- D419 - Tumeur à évolution imprévisible ou inconnue d'un organe urinaire, sans précision

Codes CIM10 leucémies aiguës

- C910 - Leucémie lymphoblastique aiguë
- C918 – Leucémies de Burkitt
- C920 - Leucémie myéloblastique aiguë
- C923 – Sarcome myéloïde
- C924 – Leucémie promyélocytaire aiguë
- C925 – Leucémie myélomonocytaire aiguë
- C926 – Leucémie myéloïde aigue avec anomalies 11q23
- C928 - Leucémie myéloïde aiguë avec dysplasie multilignées
- C930 – Leucémie myélomonocytaire aiguë
- C940 – Leucémie érythroïde aiguë
- C942 – Leucémie aigu à mégacaryocytes
- C944 – Myélofibrose aiguë
- C947 – Leucémies précisée
- C950 – Leucémie aiguë à cellules non précisées
- C835 - lymphomes lymphoblastiques
- C837 - lymphome de Burkitt.

Annexe 3 : Liste des établissements hospitaliers et leur commune d'implantation

Établissement	Ville
Centre hospitalier du Pays d'Aix	Aix-en-Provence
Clinique Axiom	Aix-en-Provence
Polyclinique du Parc Rambot	Aix-en-Provence
Centre hospitalier Joseph Imbert	Arles
Clinique Jeanne d'Arc	Arles
Centre hospitalier Edmond Garcin	Aubagne
Clinique la Casamance	Aubagne
Centre hospitalier Henri Duffaut	Avignon
Clinique Rhône-Durance	Avignon
Clinique de l'Étang de l'Olivier	Istres
Clinique de La Ciotat	La Ciotat
Clinique générale de Marignane	Marignane
Assistance Publique-Hôpitaux de Marseille	Marseille
Institut Paoli-Calmette	Marseille
Hôpital Saint-Joseph	Marseille
Hôpital Européen	Marseille
Clinique Bouchard	Marseille
Clinique Chantecler	Marseille
Clinique Bonneveine	Marseille
Hôpital privé de la Résidence du Parc	Marseille
Clinique Vert Coteau	Marseille
Hôpital privé Beauregard	Marseille
Hôpital privé Clairval	Marseille
Centre hospitalier de Martigues	Martigues
Clinique chirurgicale de Martigues	Martigues
Centre hospitalier régional de Nîmes	Nîmes
Polyclinique Kenval	Nîmes
Centre hospitalier de Salon de Provence	Salon-de-Provence
Clinique Vignoli	Salon-de-Provence

Annexe 4 : Liste des laboratoires d'anatomopathologies et laboratoires d'hématologies sollicités pour l'étude d'exhaustivité

Laboratoires d'anatomopathologies

Publics :

Laboratoire d'anatomie pathologique-neuropathologique - Hôpital de la Timone (Marseille)
Laboratoire d'anatomie pathologique - Hôpital Nord (Marseille)
Laboratoire d'anatomie pathologique de l'Institut Paoli-Calmette (Marseille)
Service d'Anatomie et Cytologie Pathologiques de l'hôpital Saint Joseph (Marseille)
Service d'Anatomie et Cytologie Pathologiques du centre hospitalier du Pays d'Aix (Aix-en-Pce)
Service d'Anatomie et Cytologie Pathologiques CHRU de Nîmes (Nîmes)

Privés :

Prado pathologie (Marseille)
SCP Chollat-Namy Calvet Patte (Marseille)
Laboratoire d'Anatomie et de cytologie pathologique Pellegrin-Martin (Marseille)
SCP Gompel Sophia et Tortel Mariann (Marseille)
SudPath (La Ciotat)
SCM Docteurs Aubert et Severy (Martigues)
Centre d'Anatomie et Cytologie Pathologiques Niel (Martigues)
L'Anapath (Eguilles)
SCP Masset-Berthier (Arles)
Cabinet d'Anatomie Pathologique et de Cytologie (Cavaillon)
Centre d'Anatomo-Cytopathologique Histosud (Le Pontet)
Laboratoire Nemaupath (Nîmes)
Cabinet d'Anatomie et Cytologie Pathologiques Valerie Laffay et Michèle Robert (Nîmes)

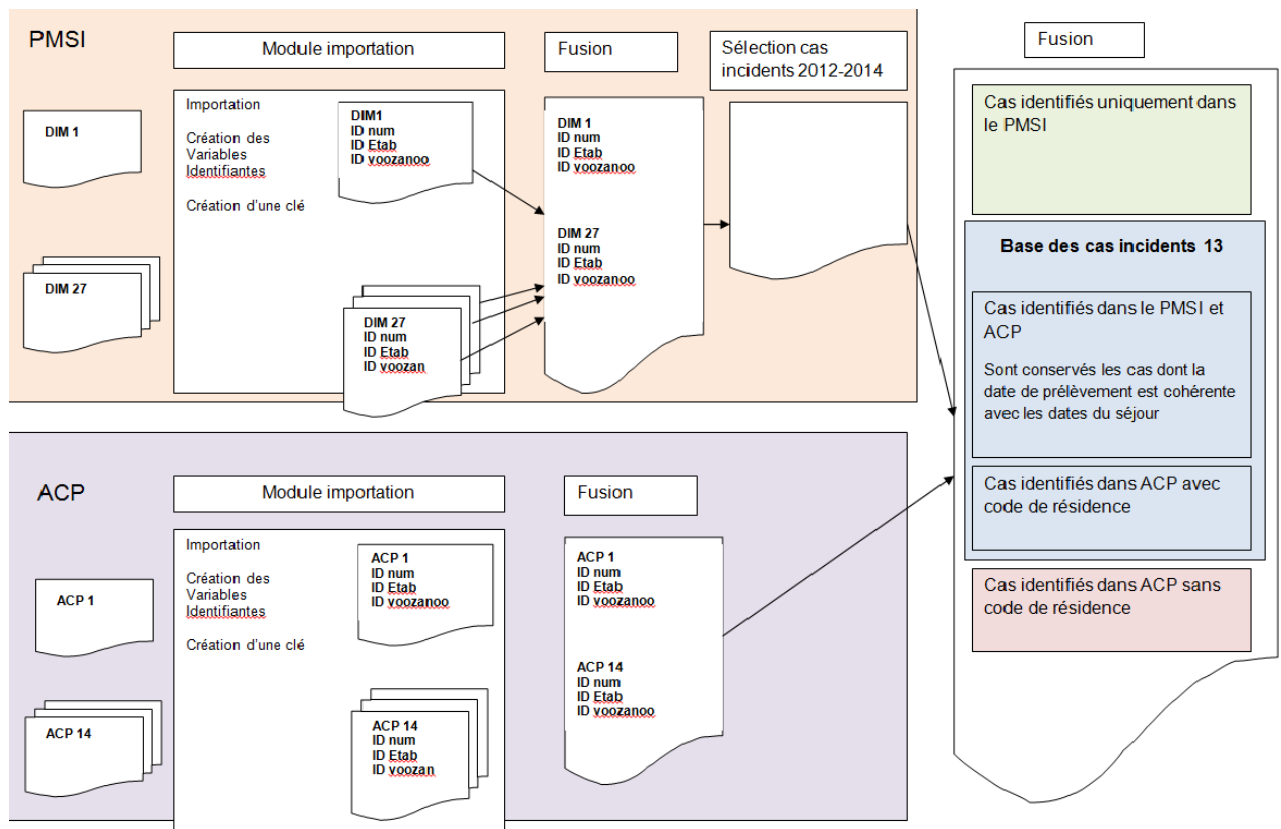
Laboratoires d'hématologie

Laboratoire d'hématologie - Hôpital de la Timone (Marseille)
Laboratoire d'hématologie de l'Institut Paoli-Calmette (Marseille)
Laboratoire d'hématologie de l'hôpital Saint Joseph (Marseille)
Laboratoire d'hématologie du Centre hospitalier du Pays d'Aix (Aix-en-Provence)
Laboratoire d'hématologie du Centre hospitalier de Martigues (Martigues)
Laboratoire d'hématologie du Centre hospitalier de Salon (Salon-de-Provence)
Laboratoire d'hématologie du Centre hospitalier Joseph Imbert (Arles)
Laboratoire d'hématologie du Centre hospitalier Henri Duffaut (Avignon)
Laboratoire d'hématologie du Centre hospitalo-universitaire de Nîmes (Nîmes)

Annexe 5 : Liste des codes Adicap

Appareil/Organe	Pathologie/ comportement	Pathologie/ comportement	Pathologie/ comportement
UC	A7A0	R7C6	V7T0
UP	A7A1	R7C8	X7E0
UT	A7A2	S7S0	X7E1
UV	A7A3	S7W0	X7E2
UR	A7A4	S7W2	X7E3
	A7A5	S7W3	X7E4
	A7C4	S7W4	X7E6
	A7C6	S7W6	X7E8
	A7D0	S7Z0	B7K0
	A7F0	S7Z1	B7K1
	A7F2	S7Z2	B7K2
	A7H6	S7Z3	B7K3
	A7H8	S7Z4	B7K4
	A7K0	S7Z6	B7K5
	A7K2	U7A0	B7K6
	A7K4	U7A1	B7K7
	A7K6	U7A2	B7K8
	A7K8	U7A3	B7K9
	A7L2	U7G0	B7L0
	A7L4	U7G1	B7L2
	A7L6	U7G2	B7L5
	A7M0	U7G3	E7T0
	A7M4	U7H0	E7T1
	A7M4	U7H1	E7T2
	A7M5	U7H2	E7T3
	A7S0	U7H3	E7T4
	A7S2	U7K0	E7T5
	A7V0	U7K1	E7T6
	A7V2	U7K2	E7T8
	A7V4	U7K3	E7T9
	A7V5	U7L0	E7V1
	A7V8	U7L1	E7V2
	A7W0	U7L2	E7V3
	A7W2	U7L3	E7V4
	A7W4	U7N0	E7V5
	A7W6	U7N3	E7V6
	A7Y1	U7P0	E7V7
	L7A0	U7P1	E7V8
	Q7A0	U7P2	D7S5
	R7C0	U7P3	F7A0
	R7C5	V7A0	F7A1
	R7C6	V7T0	F7A2

Annexe 6 : Schéma de croisement des données prévu dans le protocole



Annexe 7 - Rappel sur la méthode capture-recapture

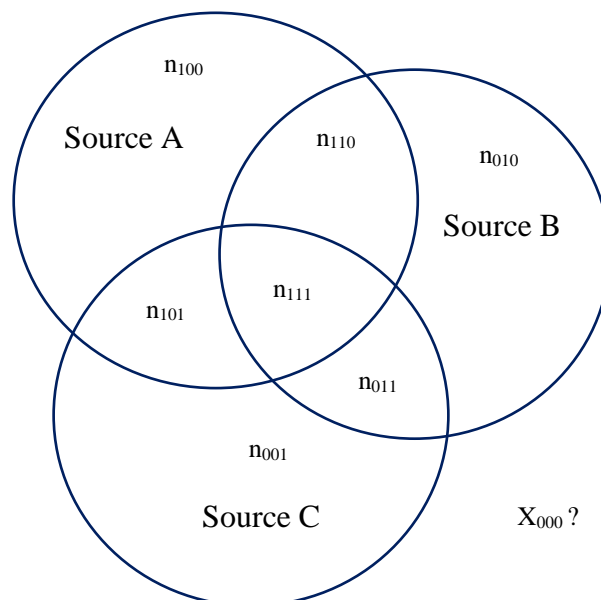
1. Principe

La méthode de capture-recapture permet, en épidémiologie, d'estimer le nombre total de cas d'une maladie donnée à partir de plusieurs sources, d'affiner la prévalence ou l'incidence à partir d'études portant sur la population générale, d'évaluer la complétude d'un système de surveillance et de fournir un intervalle de confiance sur le nombre de cas total [1-3].

Comme le montre, la Figure 1, la méthode permet, après avoir identifiés les cas communs entre ces sources, d'estimer sous certaines hypothèses, le nombre de cas présents dans aucune des sources (X_{000}).

I FIGURE 1 I

Répartition des cas d'une maladie recensés par trois sources A, B, C : Principe de la méthode capture-recapture



Les indices 1 et 0 correspondent à la présence (1) et l'absence des cas (0) dans les sources respectives A, B et C.

2. Conditions d'application

La fiabilité des résultats dépend du respect de différentes conditions d'applications.

Identification des vrais cas : l'identification de faux cas par une source induit une surestimation de nombre total de cas et une sous-estimation de l'exhaustivité pour chacune des autres sources. Par ailleurs, une définition des cas différente selon les systèmes induit également une surestimation du total (N).

Identification de tous les vrais cas communs : une surestimation des cas communs entrainera une sous-estimation du nombre total de cas et inversement. S'il n'existe pas d'identifiant commun unique entre les différents systèmes, l'identification des cas communs est basée sur une combinaison de critères.

Même période et même zone d'étude : si l'une des sources considère une zone géographique ou une période de l'étude différente de celles des autres systèmes, les estimations obtenues peuvent être biaisées.

Homogénéité de capture : tous les cas de la population étudiée ont la même probabilité d'être identifiés par une même source, même si cette probabilité peut être différente selon les sources. Dans une source, la notification des cas ne doit pas être liée à une ou plusieurs caractéristiques individuelles.

Indépendance entre les sources : cette condition implique que les cas recensés dans une source ne dépendent pas du fait qu'ils soient recensés ou non dans une autre source. Si l'identification des cas par un système augmente la probabilité de capture par un autre système, une dépendance positive sera constatée, entraînant une sous-estimation du nombre total. À l'inverse, si l'identification des cas par une source diminue la capacité de capture par une autre source, cela entraînera une surestimation causée par une dépendance négative.

Population d'étude close : aucun mouvement de population ne doit être constaté sous peine de biaiser le nombre de cas communs et donc la taille de la population.

Hook *et al* considère qu'homogénéité de capture et indépendance des sources sont liées aux autres conditions applications [1]. La plupart du temps, le non-respect des trois dernières conditions d'application peut être modélisé comme des variations de l'hypothèse d'indépendance, qui reste donc la condition première à considérer.

3. Estimation du nombre total de cas avec deux sources

L'estimation du nombre total de cas avec deux sources est possible sous l'hypothèse d'indépendance entre les sources. Le Tableau 1 présente la répartition des cas selon leur présence dans chacune des deux sources. Dans ce tableau, la présence dans une source est codée avec l'indice 1 et l'absence par l'indice 0. La cellule correspondant au nombre de cas identifiés par aucune des sources (n_{00}) est celle à estimer.

I TABLEAU 1 I

Tableau de contingence à deux sources

		Source A		Total
		Présent	Absent	
Source B	Présent	n_{11}	n_{01}	N_B
	Absent	n_{10}	n_{00}	
Total		N_A		$N = n_{11} + n_{10} + n_{01} + n_{00}$

Sous l'hypothèse d'indépendance des sources, les estimateurs de Lincoln-Petersen ou estimateur du maximum de vraisemblance (Tableau 2) permettent d'estimer n_{00} ainsi que sa variance [4]. Lorsque les effectifs sont faibles, Chapman et Wittes ont proposé d'autres estimateurs ou estimateurs presque non biaisés (Tableau 2) [5]. Un intervalle de confiance (IC) pour N se calcule dans les deux cas de la manière suivante :

$$IC\ 95\ \% \ N = \hat{N} \pm 1,96 \times \sqrt{Var(\hat{N})}$$

Le taux d'exhaustivité de la source A est $P_{A=} = \frac{N_A}{N}$ où N_A est le nombre de cas recensés dans la source A et N le nombre de cas estimés.

I TABLEAU 2 I

Formules des estimateurs du nombre total de cas, méthode capture-recapture à deux sources

Valeurs estimées		Estimateur du maximum de vraisemblance	Estimateur presque non biaisé
Cellule vide	\hat{n}_{00}	$\frac{n_{01} \times n_{10}}{n_{11}}$	$\frac{n_{01} \times n_{10}}{(n_{11} + 1)}$
Nombre de cas	\hat{N}	$\frac{N_A \times N_B}{n_{11}}$	$\left[\frac{(N_A + 1) \times (N_B + 1)}{n_{11} + 1} \right] - 1$
Variance du nombre de cas	$Var(\hat{N})$	$\frac{N_A \times N_B \times n_{01} \times n_{10}}{n_{11}^3}$	$\frac{(N_A + 1) \times (N_B + 1) \times n_{01} \times n_{10}}{(n_{11} + 1)^2 \times (n_{11} + 2)}$

4. Estimation du nombre total de cas avec plusieurs sources.

Avec trois sources, le Tableau 3 indique la répartition des cas selon leur présence ou non dans les sources.

I TABLEAU 3 I

Tableau de contingence avec trois sources

		Source A			
		Présent		Absent	
Source C	Présent	n_{111}	n_{101}	n_{011}	n_{001}
	Absent	n_{110}	n_{100}	n_{010}	n_{000}

Les indices 1 et 0 correspondent à la présence (1) et l'absence des cas (0) dans les sources respectives A, B et C

4.1. Évaluation de la dépendance avec plus de 2 sources

Deux sources sont indépendantes si la probabilité moyenne d'apparaître dans les cas communs (p_{AB}) est égale au produit des probabilités moyennes d'apparaître dans l'une des 2 sources ($p_A \times p_B$). En cas d'inégalité, il existe une dépendance négative si la probabilité moyenne est inférieure au produit des deux probabilités moyennes et une dépendance positive dans le cas contraire. Une dépendance positive des sources tend à sous-estimer le nombre total de cas et une dépendance négative des sources produira un estimateur sur estimé.

Lorsque la méthode capture-recapture est appliquée avec plus de deux sources, la dépendance entre les sources peut être évaluée et prise en compte pour l'estimation du nombre total de cas. Wittes *et al* propose d'estimer N pour chaque combinaison des sources deux à deux et d'évaluer la dépendance entre les sources en comparant les estimations. Si un couple de sources donne une estimation très différente des autres couples, une dépendance entre ces deux sources peut être suspectée [6-7].

L'indépendance entre deux sources peut être testée avec un test du Chi2 appliqué sur un tableau de contingence 2x2 à l'intérieur d'une troisième source. En outre, le calcul des rapports

de cote (OR) permet d'indiquer le sens et la force des dépendances éventuelles. Si l'OR est inférieur à 1, il existe une dépendance négative entre les sources. Si l'OR est supérieur à 1, il existe une dépendance positive entre les sources.

4.2. Estimation du nombre total de cas à l'aide des modèles log-linéaires

L'estimation du nombre total de cas avec plus de deux sources permet de prendre en compte la dépendance entre les sources. Ainsi, dans le cas d'une étude réalisée à partir de trois sources ou plus, la modélisation log-linéaire permet de prendre en compte les interactions d'ordre n-1 entre les sources (n étant le nombre de sources), l'hétérogénéité de la population, et d'avoir ainsi une meilleure estimation de la fraction non observée de la population.

Le modèle log-linéaire permet d'étudier les relations entre plusieurs variables qualitatives croisées dans un tableau de contingence. Il représente le logarithme népérien de la fréquence attendue d'une cellule du tableau comme une combinaison linéaire d'effets principaux et d'interaction. Son application dans la méthode capture-recapture est basée sur l'hypothèse d'absence d'interaction entre les 3 sources. En effet, le tableau de contingence a une cellule structurellement vide correspondant à l'absence de notification de cas dans les différentes sources. Le modèle log-linéaire utilise toutes les cellules non vides afin d'estimer les effectifs attendus, y compris celui correspondant à une absence totale de notification. La présence d'une cellule vide dans le tableau de contingence rend impossible l'ajustement du modèle en tenant compte de l'interaction d'ordre maximum entre toutes les sources.

Ainsi, le modèle saturé pour trois sources est le suivant :

$$\ln(F_{ijk}) = \theta + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

F_{ijk} correspond à la fréquence attendue pour la cellule ijk du modèle considéré. λ_i^A , λ_j^B , λ_k^C correspondent aux effets principaux de la présence dans chaque source. λ_{ij}^{AB} , λ_{ik}^{AC} , λ_{jk}^{BC} sont les termes d'interactions d'ordre 2 de la présence dans chacune des deux sources indiquées. Ici, i , j et k varient tous les trois de 0 à 1.

Avec trois sources huit modèles peuvent être construits : le modèle simple sans interactions assumant l'indépendance des sources, les modèles avec 2 sources indépendantes et le modèle avec dépendance des sources. Le nombre de modèles possible augmente rapidement avec le nombre de sources : 113 modèles pour 4 sources et 6 893 pour 5 sources, par exemple.

4.2.1. Validité et choix d'un modèle

L'estimation de l'adéquation du modèle aux données peut être fait par des méthodes non Bayésiennes, plus simples, et Bayésiennes, plus complexes non exposées ici.

Il s'agit de choisir l'estimateur dérivé du modèle jugé le plus optimal c'est-à-dire, celui qui présente la déviance (G^2) la plus faible et un test de la déviance non significatif. La sélection du modèle repose sur le principe de parcimonie, le modèle le plus simple est généralement retenu. Deux critères d'information basés sur une fonction de la déviance sont principalement utilisés, le critère d'information d'Akaike (AIC) et le critère d'information Bayésien (BIC) :

$$AIC = G^2 - 2 \times \text{ddl} \quad (\text{avec ddl : nombre de degrés de liberté})$$

$$BIC = G^2 - (\log N_{\text{obs}} / 2\pi \times \text{ddl}) \quad \text{où } N_{\text{obs}} \text{ est le nombre de cas observés}$$

Le modèle s'adaptant aux connaissances épidémiologiques et ayant les plus faibles AIC ou BIC pourra être retenu [1]. Hook *et al* suggère qu'il vaut mieux se référer au BIC qui serait plus performant dès que l'on a plus de 46 cas mais beaucoup d'études utilisent l'AIC. En cas de dépendance entre 2 sources, Hook *et al* recommande de sélectionner plutôt un modèle contenant l'interaction de ces 2 sources.

Le modèle retenu fournit une estimation du nombre de cas notifiés dans aucune source (000), puis celle de \hat{N} grâce à la relation $\hat{N}=N_{\text{obs}} + \hat{\pi}_{000}$. Les intervalles de confiance des estimateurs du nombre total de cas peuvent être calculés d'après les formules de Bishop [8] ou basés sur la déviance [9].

4.2.3 Prise en compte de l'hétérogénéité de capture

L'hétérogénéité de capture est fréquente en épidémiologie humaine. Les sources de données ou listes utilisées ont souvent à l'origine une autre finalité que la détection de cas. Le PMSI par exemple est un recueil de données à vocation médico-économique. Le processus de sélection peut varier d'une liste à l'autre et être responsable d'une variabilité des probabilités de capture.

L'effet temps (noté t) ou liste dans les modèles utilisés en épidémiologie entraîne une variabilité de la probabilité au sein des différentes occasions de capture. Certains auteurs suggèrent qu'il existe dans tous les cas une variabilité de capture à l'échelle de l'individu (notée h). Enfin, une hétérogénéité dite de comportement (notée b) induite par la première capture, la probabilité de capture étant différente après la première capture.

Cela permet de définir différents modèles prenant en compte la variation de capture: M0 (pas de source de variation, Mt, Mh, Mb et des sources de variations combinées Mth, Mbh et Mtbh).

Plusieurs estimateurs permettant de prendre en compte l'hétérogénéité des sources ont été développés. Les plus utilisés sont Chao, Poisson2 et Darroch. Le modèle de Chao estime la borne inférieure du nombre total de cas par la méthode des moments. Les estimateurs de Darroch et Poisson ajoutent une variable d'hétérogénéité dans la matrice qui correspond à une fonction des historiques de capture. Pour les modèles de Chao et Poisson 2, les probabilités de captures sont positives, celles du modèle de Darroch peuvent être négatives.

Il peut aussi exister une hétérogénéité de capture liée à certaines caractéristiques des cas qui modifient leur probabilité de capture. Cette situation est fréquente lorsque l'on s'intéresse à des cas de cancers. Ce type d'hétérogénéité est identifié par des modèles stratifiés sur la ou les variables d'intérêts de manière à analyser des sous-groupes de cas avec une probabilité de capture plus homogène. Dans notre étude, l'analyse a été stratifiée sur l'âge, le sexe et la zone de résidence des cas. Certains auteurs utilisent également un modèle de régression logistique pas à pas incluant les covariables identifiées par la stratification, d'autres un modèle de régression loglinéaire avec covariables.

Références bibliographiques

1. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev* 1995;17(2):243-64.
2. Capture-recapture and multiple-record systems estimation I: history and theoretical development. International working group for disease monitoring and forecasting. *Am J Epidemiol* 1995;142(10):1047-58.

3. Capture-recapture and multiple-record systems estimation II: applications in human diseases. International working group for disease monitoring and forecasting. *Am J Epidemiol* 1995;142(10):1059-68.
4. Sekar CC, Deming WE. On a method of estimating birth and death rates and the extent of registration. *American statistical association journal* 1949.
5. Chapman DG. Some properties of the hypergeometric distribution. *University of California* 1951;1(7):131-60.
6. Wittes J, Sidel VW. A generalization of the simple capture-recapture model with applications to epidemiological research. *J Chronic Dis* 1968;21(5):287-301.
7. Wittes JT, Colton T, Sidel VW. Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *J Chronic Dis* 1974;27(1):25-36.
8. Bishop YMM, Fienberg SE, Holland PW. Estimating the size of a closed population. *Discrete Multivariate Analysis: Theory and Practice*. The MIT Press, 1975. p. 229-56.
9. Regal RR, Hook EB. Goodness-of-fit based confidence intervals for estimates of the size of a closed population. *Stat Med* 1984;3(3):287-91.
10. Chao A. Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 1987;43:783-791.
11. Rivest LP, Baillargeon S. Applications and extensions of Chao's moment estimator for the size of a closed population. *Biometrics* 2007;63:999-1006.
12. Darroch JN, Fienberg SE, Glonek GF, Junker BW. A three-sample multiple-recapture approach to census population estimation with heterogeneous catchability. *J Am Stat Ass* 1993;88(423):1137-47.
13. Plante N, Rivest LP, Tremblay G. Stratified capture-recapture estimation of the size of a closed population. *Biometrics* 1998;54:47-60.
14. Baillargeon S, Rivest LP. Rcapture: Loglinear models for capture-recapture in R. *J Stat Soft* 2007;19(5):1-31.
Disponible à partir de l'URL : https://www.google.fr/search?q=Baillargeon+S,+Rivest+LP.+Rcapture:+Loglinear+model+s+for+capture-recapture+in+R.+J+Stat+Soft+2007:19%285%29:&ie=utf-8&oe=utf-8&qws_rd=cr&dcr=0&ei=zMO7We-dNIWzaZnnucqO
15. R software : Package Rcapture. Disponible à partir de l'URL : <https://cran.r-project.org/web/packages/Rcapture/Rcapture.pdf>
16. Bailly L. Validation et exploitation d'un registre histologique des cancers. Estimation par capture-recapture de l'exhaustivité par modèles log-linéaires et selon les modèles écologiques Mtbh bayésiens. *Biostatistique épidémiologie*. Université Montpellier I, 2011. Français.
17. Berger F, Parent du Châtelet I, Bernillon P et Gallay A. Surveillance des infections invasives à méningocoque en France métropolitaine en 2005 – Évaluation quantitative par la méthode de capture-recapture à trois sources. Saint-Maurice (France) : Institut de veille sanitaire, août 2010, 43 p. Disponible à partir de l'URL : www.SpFrance.sante.fr

Annexe 8 - R-capture

L'analyse de R-Capture dans ce contexte consiste dans un premier temps à réaliser une matrice agrégée de l'historique de capture pour chaque unité « capturée », soit chaque cas de cancer dans notre étude. Chaque ligne de la matrice correspond à un historique de capture et chaque colonne à une source potentielle de capture, la dernière colonne donnant le nombre de cas agrégé répondant à chaque historique de capture possible comme illustrée dans le Tableau 1.

À partir de cette matrice, douze modèles sont ensuite ajustés, ils correspondent aux modèles de détection homogène (M0), de détection hétérogène avec les estimateurs Chao, Poisson2, Gamma 3.5 ou Darroch (Mh), des modèles en fonction du temps ou liste (Mt), ou d'autres modèles (Mb) qui attribue une probabilité différente aux individus après leur première capture (fixée aléatoirement ici) et ceux combinant plusieurs caractéristiques (Mtb et Mth).

I TABLEAU 1 I

Exemple de matrice agrégée de l'historique de capture par les 3 sources de l'étude : Revela13, PMSI et ACP

Cas Revela13	Cas ACP	Cas PMSI	Nombre de cas
1	0	0	-
0	1	0	-
1	1	0	-
0	0	1	-
1	0	1	-
0	1	1	-
1	1	1	-

Les indices 1 et 0 correspondent à la présence (1) et l'absence des cas (0) respectivement dans les trois sources citées

Évaluation de l'hétérogénéité de capture

La fonction *descriptive* permet de décrire les fréquences de capture-recapture suivantes :

- f_i le nombre de cas capturé par i fois
- u_i le nombre de cas capturé la première fois par la liste i
- v_i le nombre de cas capturé la dernière fois par la liste i
- n_i le nombre de cas capturé par la liste i

Si le nombre n_i varie pour chacune des listes, il existe un effet lié à la source.

L'hétérogénéité de capture des cas peut être évaluée à partir des graphes des $\log\left(\frac{f_i}{\binom{t}{i}}\right)$ et des $\log(u_i)$ en fonction des i listes.

Une courbe convexe ou pas complètement linéaire est en faveur d'une hétérogénéité de capture.

Choix du modèle le mieux ajusté

La fonction *closedep* permet de produire les douze modèles de capture-recapture décrit précédemment avec les différents estimateurs. Le choix du modèle le mieux ajusté est empirique et se fait à partir de plusieurs critères :

- Minimisation de la déviance jugée à partir des critères d'information AIC et le BIC. Pour les modèles basés sur l'estimateur de Chao, une faible déviance signifie la présence d'une hétérogénéité mais pas que la limite minimale estimée est non biaisée.
- Précision de l'estimation de N

Deux outils sont disponibles dans le package *Rcapture* pour aider au choix du modèle :

- La fonction *boxplot.closedep* permet de construire le graphe des boxplot des résidus de Pearson, le modèle le mieux ajusté étant celui avec les résidus les plus proches de zéro.
- La fonction *uifit* permet de comparer les valeurs observées et les valeurs prédites par chaque modèle permettant de choisir celui qui produit les valeurs prédites les plus proches des valeurs observées.

Le modèle retenu à cette étape est introduit dans la fonction *closedpMS.t* qui permet de tester les interactions entre les sources ou listes. Le choix du modèle est fait sur les critères suivants :

- Respect du critère de parcimonie
- Prise en compte de la dépendance des sources en retenant le modèle avec l'interaction correspondante
- Minimisation de la déviance jugée à partir des critères d'information AIC et le BIC. Pour les modèles basés sur l'estimateur de Chao, une faible déviance signifie la présence d'une hétérogénéité mais pas que la limite minimale estimée est non biaisée.
- Précision de l'estimation de N

Estimation du nombre total de cas

La fonction *closedpCI.t* produit un estimateur du nombre de cas total et son erreur standard. Deux types d'intervalles de confiance peuvent être calculés :

- un intervalle de confiance symétrique assumant la normalité de la distribution ;
- un intervalle de confiance basé sur la valeur de N maximisant la vraisemblance multinomiale, il est asymétrique. La fonction *profileCI* permet de faire un graphe de distribution des profils de vraisemblance.

Annexe 9 : Analyse de sensibilité avec l'estimateur de Chao

Estimation de l'exhaustivité et intervalle de confiance à 95% de chaque source pour les tumeurs urothéliales selon le modèle log-linéaire de Chao, Bouches-du-Rhône, 2013-2014

Source	Exhaustivité (%)	IC95%
Revela13	84,2	83,8 – 84,5
PMSI	91,7	91,3 – 92,1
ACP	87,0	86,6 – 87,4
3 sources	98,0	96,2 – 99,9

Estimation de l'exhaustivité et intervalle de confiance à 95% de chaque source pour les cancers du rein selon le modèle log-linéaire de Chao, Bouches-du-Rhône, 2013-2014

Source	Exhaustivité (%)	IC95%
Revela13	84,7	83,7 – 85,7
PMSI	87,7	86,7 – 88,8
ACP	81,7	80,7 – 82,7
3 sources	98,6	97,5 – 99,8

Annexe 10 : Analyse stratifiée des tumeurs urothéliales

Nombre de tumeurs urothéliales et du taux d'exhaustivité stratifiés en fonction de la classe d'âge et la zone de résidence et du sexe, Bouches-du-Rhône, 2013-2014

Secteur de résidence	Sexe	Classe d'âge	Nombre de cas signalés			Cas inclus total	N (95% IC)	Exhaustivité du système		
			REVELA13	ACP	DIM			REVELA13	ACP	DIM
MARSEILLE	Hommes	moins de 65 ans	89	92	95	100	100,6 (98,4 - 102,8)	88,5 (86,6 - 90,4)	91,5 (89,5 - 93,5)	94,4 (92,4 - 96,5)
		65 - 85 ans	257	266	277	295	295,6 (293,8 - 297,4)	86,9 (86,4 - 87,5)	90,0 (89,4 - 90,5)	93,7 (93,1 - 94,3)
		plus de 85 ans	45	49	51	53	53,0 (52,6 - 53,4)	84,9 (84,3 - 85,6)	92,4 (91,8 - 93,2)	96,2 (95,5 - 96,9)
	Femmes	moins de 65 ans	34	30	36	38	38,0 (37,6 - 38,4)	89,5 (88,5 - 90,4)	78,9 (78,1 - 79,8)	94,7 (93,8 - 95,7)
		65 - 85 ans	63	66	61	72	72,0 (71,8 - 72,2)	87,5 (87,3 - 87,7)	91,6 (91,4 - 91,9)	84,7 (84,5 - 85,0)
		plus de 85 ans	22	24	23	25	25,0 (24,8 - 25,2)	88,0 (87,3 - 88,7)	96,0 (95,2 - 96,8)	92,0 (91,3 - 92,7)
HORS MARSEILLE	Hommes	moins de 65 ans	114	105	122	127	128,5 (124,7 - 132,2)	88,7 (86,2 - 91,4)	81,7 (79,4 - 84,2)	94,9 (92,3 - 97,8)
		65 - 85 ans	267	274	291	326	331,8 (324,3 - 339,2)	80,5 (78,7 - 82,3)	82,6 (80,8 - 84,5)	87,7 (85,8 - 89,7)
		plus de 85 ans	55	63	64	72	75,7 (67,7 - 83,7)	72,7 (65,7 - 81,2)	83,2 (75,3 - 93,1)	84,5 (76,5 - 94,5)
	Femmes	moins de 65 ans	26	25	26	28	28,1 (27,3 - 28,9)	92,5 (90,0 - 95,2)	89,0 (86,5 - 91,6)	92,5 (90,0 - 95,2)
		65 - 85 ans	56	61	63	72	73,7 (69,6 - 77,8)	75,9 (72,0 - 80,5)	82,8 (78,4 - 87,6)	85,5 (81,0 - 90,5)
		plus de 85 ans	19	27	29	30	30,5 (28,1 - 32,9)	62,3 (57,7 - 67,6)	88,5 (82,1 - 96,1)	95,1 (88,1 - 103,2)