



## Le Comité de surveillance Statistique

### Délibération STAT n° 10/2017 du 9 mai 2017

**Objet** : Demande formulée par la Fondation Registre du Cancer et le groupe de recherche Interface Demography de la VUB afin d'obtenir de la Direction générale Statistique – Statistics Belgium la communication de données d'étude codées (recensement 2001 couplé aux causes de décès 2001-2011) en vue de la réalisation d'une étude sur la relation entre le statut socio-économique ou la migration et l'incidence, la survie et la mortalité du cancer (STAT-MA-2017-009)

Le Comité de Surveillance Statistique (ci-après "le Comité") ;

Vu la loi du 4 juillet 1962 *relative à la statistique publique* (ci-après "la loi statistique publique") ;

Vu la loi du 8 décembre 1992 *relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel* (ci-après "la LVP") ;

Vu l'arrêté royal du 13 février 2001 *portant exécution de la loi du 8 décembre 1992 relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel* (ci-après "l'arrêté royal du 13 février 2001") ;

Vu l'arrêté royal du 7 juin 2007 *fixant les modalités relatives à la composition et au fonctionnement du Comité de surveillance Statistique institué au sein de la Commission de la protection de la vie privée* ;

Vu la demande de la Fondation Registre du Cancer et du groupe de recherche Interface Demography de la VUB, reçue le 14/02/2017; Vu les documents complémentaires reçus le 02/05/2017 et le 08/05/2017 ;

Vu la demande d'avis technique et juridique adressée au Service public fédéral Économie, PME, Classes moyennes et Énergie (Direction générale Statistique – Statistics Belgium) le 29/03/2017 ;

Vu l'avis technique et juridique reçu le 25/04/2017 ;

Vu le rapport du Président ;

Émet, après délibération, la décision suivante, le 9 mai 2017 :

## **I. OBJET DE LA DEMANDE**

1. La demande vise à ce que la Fondation Registre du Cancer et le groupe de recherche Interface Demography de la VUB, ci-après les Chercheurs, soient autorisés à recevoir de la Direction générale Statistique – Statistics Belgium (ci-après la DGSSB) la communication de données d'étude codées (données de l'enquête socio-économique (recensement) de 2001 et causes de décès de 2001-2011 de la Banque nationale de données de mortalité) pour les coupler ensuite avec les caractéristiques d'incidence et de tumeur de la Fondation Registre du Cancer, et ce en vue de la réalisation d'une étude sur la relation entre le statut socio-économique ou la migration et l'incidence, la survie et la mortalité du cancer.

2. La demande vise également à ce que soit approuvé le contrat de confidentialité à conclure entre la DGSSB et les Chercheurs à la suite de cette communication.

3. Dans la délibération n° 16/086 du 20 septembre 2016, la section Santé du Comité sectoriel de la Sécurité Sociale et de la Santé a déjà octroyé aux Chercheurs une autorisation en vue du couplage (avec l'intervention de la Plateforme eHealth) de données à caractère personnel codées relatives à la santé de la Fondation Registre du Cancer avec des données de la DGSSB et de la communication de ces données dans le cadre de l'étude scientifique précitée, moyennant certes une autorisation du Comité de surveillance Statistique pour la communication des données d'étude codées provenant de la DGSSB.

## **II. EXAMEN DE LA DEMANDE**

### **A. LÉGISLATION APPLICABLE**

#### **A.1. Loi statistique publique**

4. Sur la base des articles 15 et 15*bis* de la loi statistique publique, la DGSSB est habilitée, après autorisation du Comité de surveillance statistique et moyennant un contrat de confidentialité approuvé

par ce même Comité, à communiquer des données d'étude codées aux destinataires mentionnés dans la loi statistique publique aux conditions fixées dans cette même loi.

## **A.2. LVP et arrêté royal du 13 février 2001**

5. En vertu de l'article 1, § 1 de la LVP et de l'article 1, 3° de l'arrêté royal du 13 février 2001, des données d'étude codées relatives à des personnes physiques identifiées ou identifiables constituent des données à caractère personnel dont le traitement n'est autorisé qu'aux conditions fixées dans la LVP et dans l'arrêté royal du 13 février 2001.

## **B. BASE JURIDIQUE**

6. Les Chercheurs font partie des destinataires de données énumérés dans la loi statistique publique, plus précisément au sens de l'article article 15, premier alinéa, 1° de la loi statistique publique en ce qui concerne la Fondation Registre du Cancer<sup>1</sup> et au sens de l'article 15, premier alinéa, 4° de la loi statistique publique en ce qui concerne le groupe de recherche Interface Demography de la VUB<sup>2</sup>.

7. Les Chercheurs entrent donc en principe en ligne de compte pour être autorisés à recevoir les données demandées.

## **C. FINALITÉ**

8. Les données à caractère personnel doivent être collectées pour des finalités déterminées, explicites et légitimes (article 4, § 1, 2° de la LVP).

9. Les Chercheurs souhaitent réaliser une étude scientifique sur la relation entre le statut socio-économique ou la migration et l'incidence, la survie et la mortalité du cancer (les différentes questions de l'étude sont énoncées dans la demande).

Une première finalité de l'étude est de vérifier comment des différences individuelles dans des facteurs socio-économiques se traduisent par des différences dans l'incidence du cancer. Outre le sexe et l'âge, le contexte socio-économique est également important pour pouvoir étudier adéquatement des schémas dans l'incidence du cancer. La prise en compte du statut socio-économique permet de diviser

---

<sup>1</sup> En vertu de l'article 45 *quinquies* de l'arrêté royal n° 78 du 10 novembre 1967 *relatif à l'exercice des professions des soins de santé*, la Fondation Registre du Cancer, en tant que fondation d'utilité publique, est notamment chargée de "1° établir des rapports concernant l'incidence des différentes formes de cancer, ainsi que sa prévalence et la survie des patients ; 2° réaliser des études (contrôle de cas et étude de cohortes) sur les causes du cancer ; 3° effectuer une analyse de la répartition géographique des différentes formes de cancer, son incidence, sa tendance et ses conséquences afin de pouvoir examiner les causes possibles et de pouvoir comparer les facteurs de risques ; ...".

<sup>2</sup> Le groupe de recherche Interface Demography de la VUB a déjà été autorisé précédemment par le Comité de surveillance Statistique à obtenir des données d'étude codées de la DGSSB, et ce par la délibération STAT n° 20/2014.

la population en strates avec davantage de précision ainsi que de mieux prédire des différences d'incidence. Cette finalité sera principalement développée par les chercheurs de la Fondation Registre du Cancer.

Une deuxième finalité consiste à établir, pour la population belge, la carte de la répartition sociale en termes d'incidence, de survie et de mortalité pour les différents types de cancer. En étudiant les inégalités socio-économiques dans ces paramètres, l'étude pourra identifier les sites de cancer présentant les plus grandes inégalités et des différences en matière de mortalité due au cancer pourront être analysées selon les différences d'incidence du cancer et de survie. Cette finalité sera principalement développée par le groupe de recherche de la VUB.

10. Afin de pouvoir réaliser l'étude, il est nécessaire de pouvoir coupler les chiffres de mortalité et les causes de décès de la Banque nationale de données de mortalité ainsi que les données de l'enquête socio-économique (recensement) de 2001 avec des caractéristiques d'incidence et de tumeur de la Fondation Registre du Cancer<sup>3</sup>.

11. Les finalités de l'étude répondent aux exigences précitées de la LVP.

12. En vertu de l'article 15 de la loi statistique publique, les données à caractère personnel codées doivent être collectées à des fins statistiques ou scientifiques. En l'occurrence, les données sont destinées à une étude scientifique relative à la relation entre le statut socio-économique ou la migration et l'incidence, la survie et la mortalité du cancer.

13. D'après l'institution de gestion, la demande ne pose aucun problème sur le plan de la finalité ; la finalité statistique est garantie. Le Comité adhère à cet avis.

#### **D. DONNÉES**

14. Les Chercheurs décrivent en détail dans leur demande les variables qu'ils souhaitent recevoir de la DGSSB, au départ de la banque de données couplée de l'enquête socio-économique (recensement) de 2001 et de la Banque nationale de données de mortalité (2001-2011) :

- NISS (en vue du couplage avec les données de la Fondation Registre du Cancer)
- année de naissance
- âge au moment du recensement
- intervalle de temps entre la date de migration et la date du recensement
- niveau de formation

---

<sup>3</sup> La population d'étude comporte toutes les personnes qui, au moment du recensement de 2001, séjournaient légalement en Belgique et étaient alors âgées de 25 ans ou plus. Seule la première incidence d'un des types de tumeur sélectionnés est prise en compte.

- niveau de confort du logement
- revenus du ménage
- propriété du logement
- situation étudiant, travailleur, pensionné, ...
- statut professionnel
- nationalité d'origine
- secteur d'activité de l'organisme dans lequel la personne concernée a travaillé/travaille
- nationalité actuelle
- pays de naissance
- nombre d'années en Belgique
- nombre d'enfants et âge de la mère au premier enfant
- état civil
- composition du ménage (lipro)
- santé subjective – malade de longue durée
- décès spécifique à une cause (2001-2013) – types de tumeurs sélectionnés : cancer de l'estomac, du colon, du rectum, du foie, de la tête et du cou, du pancréas, des poumons, mélanome malin, cancer du sein, du col de l'utérus, de la thyroïde, lymphome hodgkinien.

15. Le demandeur explique que le couplage des données précitées de la DGSSB aux données de la Fondation Registre du Cancer sera réalisé par l'intermédiaire de la plateforme eHealth.

Le Comité précise à cet égard que la plateforme eHealth interviendra en tant que sous-traitant de la DGSSB dans son rôle d'organisation intermédiaire, dès lors que la loi statistique publique ne lui permet pas encore de communiquer des données à caractère personnel non codées à d'autres organisations intermédiaires<sup>4</sup>. La DGSSB et la plateforme eHealth prendront les dispositions contractuelles nécessaires en application de l'article 16 de la LVP.

Le Comité précise par ailleurs que l'activité de codage et de couplage dans le chef de la plateforme eHealth doit se limiter à l'établissement d'un tableau de concordance entre l'INSS et l'INSS codé, sans réunir à cette occasion des données concrètes (dont des données sensibles relatives à la santé). Les fournisseurs de données pourront alors mettre à disposition des Chercheurs des fichiers codés qui peuvent être couplés.

16. L'institution de gestion indique dans son avis que les données sur la cause de décès initiale sont disponibles de 2001 à 2014 et qu'elles peuvent être fournies à un niveau à 4 chiffres (le niveau demandé le plus détaillé pour certaines causes de décès). Les Chercheurs sont responsables de la

---

<sup>4</sup> Ce par analogie avec les recommandations que la Commission de la protection de la vie privée a émises dans ce domaine à l'égard de la BCSS dans son rôle d'instance en charge du couplage et du codage notamment de données provenant de la DGSSB (recommandation n° 01/2011 du 9 février 2011, modifiée par la recommandation n° 02/2015 du 25 février 2015).

composition des groupes d'autres causes de décès. La fourniture de la cause de décès initiale ne devrait pas poser de problème selon la DGSSB, étant donné qu'une seule variable est demandée.

Les variables provenant du recensement de 2001 ne posent pas de problème non plus, avec quelques réserves pour la variable relative à l'origine (natmax\_9101\_svr\_recode), qui requiert probablement un travail de développement et/ou de validation dans le chef de la DGSSB.

## **E. PROPORTIONNALITÉ**

### **E.1. Quant à la nécessité d'obtenir des données codées**

17. Les Chercheurs ne peuvent recevoir les données à caractère personnel codées que si un traitement de données anonymes ne permettait pas de réaliser les finalités statistiques ou scientifiques visées (article 4 de la LVP).

18. L'étude visée ici ne peut pas supporter la "perte d'informations" d'une éventuelle anonymisation par la DGSSB (par exemple, en reprenant les données demandées dans des tableaux indiquant des totaux). La nécessité de pouvoir disposer de données non agrégées ressort clairement de la quantité et de la richesse des variables, aussi au niveau individuel, et le Comité reconnaît dès lors le besoin des données à caractère personnel codées qui sont demandées pour les finalités de recherche visées. Une communication d'informations purement anonymes ne peut ici suffire.

19. Les finalités justifient donc le traitement de données à caractère personnel codées. Vu la quantité et la richesse des variables, les Chercheurs doivent toutefois s'abstenir d'effectuer des opérations visant à convertir des données codées en données non codées.

### **E.2. Quant à la quantité de données**

20. Les données demandées sont une sélection de variables provenant de la banque de données liée de l'enquête socio-économique (recensement) de 2001 et de la Banque nationale de données de mortalité (2001-2011), dont les Chercheurs expliquent en détail dans leur demande la pertinence à l'égard des différentes finalités.

21. L'institution de gestion estime également que les données demandées (tant les causes de décès que le volet recensement et fichiers démographiques) sont proportionnelles aux questions de l'étude.

22. D'après la DGSSB, le principe de proportionnalité est donc pris en compte vu que les données demandées sont adéquates, pertinentes et non excessives à l'égard des finalités de l'étude clairement décrites. Le Comité adhère à ce point de vue et en conclut que l'article 4, § 1, 3° de la LVP est respecté.

### **E.3. Quant au délai de conservation des données**

23. Les données à caractère personnel peuvent être conservées pendant une durée n'excédant pas celle nécessaire à la réalisation des finalités pour lesquelles elles sont obtenues ou pour lesquelles elles sont traitées ultérieurement (article 4, § 1, 5° de la LVP).

24. D'après les Chercheurs, le projet d'étude durera jusqu'à fin 2018 ; ils souhaitent un accès sécurisé à l'ensemble de données codées (voir note de bas de page 4) jusqu'à 2 ans après la clôture de l'étude, à savoir jusqu'au 31 décembre 2020.

Ce délai de conservation jusqu'à 2 ans après la clôture de l'étude se justifie d'une part par la multitude et l'ampleur des questions de l'étude et d'autre part par le projet de rédiger diverses publications scientifiques sur la base des résultats de l'étude. Il est donc nécessaire de disposer encore des données après la clôture de l'étude à des fins de contrôle et de révision par des pairs pour les articles transmis. L'institution de gestion estime également qu'il s'agit d'un délai de conservation raisonnable compte tenu des finalités et du nombre de données disponibles. Le Comité en prend acte.

25. Une fois passé le délai de conservation, les données et sauvegardes doivent être complètement détruites par les Chercheurs. Il n'est pas permis de continuer à utiliser plus longtemps les données d'étude (codées) pour les mêmes finalités, sauf prolongation accordée. Si les finalités sont atteintes avant l'échéance de ce délai, les données et sauvegardes doivent être détruites par les Chercheurs avant ce terme, c'est-à-dire immédiatement après la réalisation des finalités.

## **F. SÉCURITÉ**

### **F.1 Conseiller en sécurité de l'information**

26. D'après les documents transmis par les Chercheurs, il apparaît que ces derniers disposent d'un conseiller en sécurité de l'information dont l'identité a été communiquée. Le Comité en prend acte.

### **F.2. Politique de sécurité de l'information**

27. D'après les déclarations de conformité relatives au système de sécurité de l'information, pour autant qu'elles aient été complétées conformément à la vérité, on peut établir que les 14 mesures de sécurité généralement recommandées par la Commission lors du traitement de données à caractère personnel sont réalisées. Le Comité en prend acte.

28. Les Chercheurs signalent que le traitement des données, dont la plupart sont des données relatives à la santé, se fera sous la surveillance d'un médecin.

29. Les informations fournies dans le dossier de demande et en particulier dans les déclarations de conformité relatives au système de sécurité de l'information indiquent un niveau raisonnable de sécurité (technique/ICT), d'après l'institution de gestion, qui émet un avis positif à cet égard. Le Comité en prend acte.

### **F.3. Personne physique responsable**

30. L'identité de la personne physique responsable a été communiquée. Celle-ci est personnellement responsable du respect de toutes les obligations concernant l'exécution de la loi statistique publique, de la LVP, de leurs arrêtés d'exécution, de toute autre disposition légale ou réglementaire visant la protection de la vie privée, des dispositions de la présente décision du Comité et des dispositions du contrat de confidentialité.

31. Cette personne exercera un contrôle effectif de l'utilisation licite des données fournies.

32. Les mesures dont il est question aux points F.1. à F.3. inclus, qui doivent garantir la protection et la sécurité des données d'étude transmises, comme l'exigent l'article 16 de la LVP et l'article 15*bis* de la loi statistique publique, sont efficaces si elles sont contrôlées et suivies de façon stricte dans la pratique.

### **F.4. Séparation des autres traitements**

33. Les Chercheurs doivent séparer le présent traitement des données dont il est question ici pour les finalités indiquées des autres traitements de données à caractère personnel qu'ils contrôlent éventuellement.

### **F.5. Interdiction de décodage**

34. Les Chercheurs doivent s'engager contractuellement à mettre en œuvre tous les moyens possibles afin d'éviter que soit retrouvée l'identité des personnes auxquelles les données à caractère personnel communiquées se rapportent.



#### **F.6. Interdiction de couplage**

35. Les Chercheurs ne peuvent pas tenter de coupler les données à caractère personnel obtenues à des données à caractère personnel qui leur ont déjà été transmises en application d'autres autorisations.

#### **F.7. Confidentialité**

36. Les Chercheurs s'engagent à respecter la confidentialité des données d'étude et à veiller à ce que celles-ci ne soient utilisées que par les membres de leur propre personnel en vue de l'exécution de l'étude visée.

#### **F.8. Personnes qui utilisent les données d'étude et liste de ces personnes**

37. Plusieurs données de l'ensemble de données demandé peuvent être considérées comme des données à caractère personnel au sens de l'article 7 de la LVP.

38. Vu le caractère sensible de ces données, les Chercheurs doivent respecter le chapitre III de l'arrêté royal du 13 février 2001, à savoir :

- établir une liste qui mentionne les (catégories de) personnes qui utiliseront les données communiquées. Cette liste sera constamment actualisée et tenue à la disposition de la Commission/du Comité.
- ces personnes doivent être tenues au respect du caractère confidentiel des données concernées par une obligation légale ou statutaire, ou par une disposition contractuelle équivalente. Ces personnes signeront au moins une déclaration par laquelle elles s'engagent à préserver la sécurité et le caractère confidentiel des informations.

### **G. AUTRES CONDITIONS D'UTILISATION**

#### **G.1. Diffusion des résultats**

39. Les Chercheurs doivent en effet veiller à ce qu'après l'analyse et l'utilisation des données, les résultats publiés restent anonymes et globaux de sorte que les données individuelles ne puissent pas être identifiées directement ou indirectement à partir de ces résultats.

40. Les résultats ne peuvent donc être diffusés que de manière globale et anonyme.

41. Au minimum deux semaines avant la diffusion, les Chercheurs doivent d'ailleurs soumettre la publication envisagée à la DGSSB.

## **G.2. Finalité scientifique, normes scientifiques et méthodes d'analyse**

42. Pour le Chercheur (groupe de recherche Interface Demography de la VUB), qui est destinataire au sens de l'article 15, premier alinéa, 4° de la loi statistique publique, se pose l'exigence de décrire les méthodes d'analyse qui seront utilisées lors de la recherche et de prouver que la recherche répond aux normes scientifiques en vigueur.

43. Le Chercheur en question explique ce qui suit au sujet des méthodes d'analyse employées : *"Pour analyser l'incidence et la mortalité, on utilisera aussi bien la standardisation (in)directe que la régression Poisson log-linéaire. La survie du cancer en fonction de la PSE (position socio-économique) et de l'origine sera étudiée à l'aide de techniques de pointe de survie générale, de survie spécifique du cancer et de survie relative."* [Traduction libre effectuée par le Secrétariat de la Commission vie privée, en l'absence de traduction officielle].

En ce qui concerne les normes scientifiques en vigueur, le Chercheur indique ce qui suit : *"Le projet recourt à des techniques de pointe pour analyser les données et répond à un besoin scientifique fondamental et axé sur la pratique. Le projet a également été récompensé récemment par le Fonds de la Recherche scientifique de Flandre (numéro de projet G043517N)."* [Traduction libre effectuée par le Secrétariat de la Commission vie privée, en l'absence de traduction officielle].

44. L'institution de gestion confirme dans son avis que les Chercheurs disposent des meilleurs outils et méthodes pour garantir la valeur scientifique de l'étude. Elle souligne également l'importance fondamentale de la recherche scientifique visée et le fait que le couplage des causes de décès et du recensement ouvre à cet égard des perspectives très intéressantes que l'étude entend exploiter.

45. Pour autant que le Comité ait pu en juger, le caractère scientifique tant de la finalité que de la méthodologie du projet de recherche ne peut être contesté.

## **G.3. Contrôle**

45. Les Chercheurs acceptent expressément que des représentants du Comité aient, à chaque instant et sans mise en demeure préalable, accès aux locaux et à l'infrastructure informatique où les données communiquées sont conservées, pour contrôler l'exécution des dispositions de la décision du Comité, des dispositions de la loi statistique publique et de ses arrêtés d'exécution ainsi que des dispositions du contrat de confidentialité.

47. Sur simple demande, le Comité peut obtenir l'accès à d'autres locaux et à d'autres systèmes ICT afin de contrôler si aucune violation des dispositions de sa décision, des dispositions de la loi statistique publique et de ses arrêtés d'exécution et des dispositions du contrat de confidentialité n'est commise.

## **H. LE CONTRAT DE CONFIDENTIALITÉ**

48. Les données d'étude sont communiquées aux Chercheurs en vertu d'un contrat de confidentialité conclu entre la DGSSB et les Chercheurs.

49. Le contrat de confidentialité, joint en annexe à la demande de communication des données, fixe les conditions auxquelles les données peuvent être transmises par la DGSSB et utilisées par les Chercheurs.

L'institution de gestion précise dans son avis qu'un contrat de confidentialité sera conclu avec chacun des deux Chercheurs.

50. Le contrat de confidentialité contient au moins les mentions légalement obligatoires telles que définies à l'article 15*bis* de la loi statistique publique, dont la durée du contrat de confidentialité qui, aux yeux du Comité, est conclu jusqu'au 31 décembre 2020. Cela ne signifie aucunement qu'au terme de ce délai contractuel, la confidentialité des données elles mêmes puisse être rompue. Elle doit dès lors être respectée de manière illimitée dans le temps.

51. Les dispositions contractuelles relatives à la vie privée et à la confidentialité figurant dans le contrat de confidentialité sont reprises dans la présente décision du Comité, ce qui permet ainsi également à des personnes étrangères au contrat de confidentialité de s'adresser au Comité qui peut donc contrôler le respect des conditions auxquelles les données peuvent être utilisées par les Chercheurs.

## **III. DÉCISION GÉNÉRALE**

52. Lors du traitement des données à caractère personnel obtenues, les Chercheurs doivent tenir compte de la LVP, de la loi statistique publique, de leurs arrêtés d'exécution et de toute autre disposition légale ou réglementaire de protection de la vie privée, des dispositions de la présente décision du Comité et des dispositions du contrat de confidentialité qu'ils auront conclu avec la DGSSB.

## **IV. DÉCISION SPÉCIFIQUE**

53. Le Comité estime que :

- les Chercheurs disposent d'un fondement légal pour réclamer les données d'étude codées qui sont demandées ;
- la communication par la DGSSB des données d'étude codées demandées aux Chercheurs est autorisée en vue des finalités précitées de recherche statistique/scientifique ;
- ces données sont livrables en tenant compte du niveau de détail de la demande et compte tenu de la réserve exprimée au point 16 ;
- le codage et le couplage de données doivent se faire conformément aux recommandations émises par la Commission de la protection de la vie privée dans ce domaine (recommandation n° 01/2011 du 9 février 2011, modifiée par la recommandation n° 02/2015 du 25 février 2015), en vertu desquelles la plateforme eHealth intervient en tant que sous-traitant de la DGSSB dans son rôle d'organisation intermédiaire, l'activité de codage et de couplage dans le chef de la plateforme eHealth devant se limiter à l'établissement d'un tableau de concordance entre l'INSS et l'INSS codé, sans réunir à cette occasion des données concrètes (dont des données sensibles relatives à la santé) ;
- la durée de conservation des données est limitée à 2 ans après la clôture du projet d'étude, à savoir jusqu'au 31 décembre 2020.

## **PAR CES MOTIFS,**

le Comité

**1° autorise** la DGSSB à communiquer les données à caractère personnel précitées à la Fondation Registre du Cancer et au groupe de recherche Interface Demography de la VUB ;

**2° décide** que la présente autorisation est valable si et aussi longtemps que les conditions de la présente délibération sont respectées ;

**3° approuve** le contrat de confidentialité y afférent, aux conditions précitées.

L'Administrateur f.f.,

Le Président,

(sé) An Machtens

(sé) Gert Vermeulen