

<p>Comité de sécurité de l'information Chambre sécurité sociale et santé</p>
--

CSI/CSSS/22/276

DÉLIBÉRATION N° 22/150 DU 7 JUIN 2022 RELATIVE À LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL RELATIVES À LA SANTÉ PSEUDONYMISÉES PAR L'AGENCE INTERMUTUALISTE ET DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES ISSUES DU DATAWAREHOUSE MARCHÉ DU TRAVAIL ET PROTECTION SOCIALE PAR LA BCSS À L'UNIVERSITÉ LIBRE DE BRUXELLES (DULBEA) DANS LE CADRE D'UNE ÉTUDE SCIENTIFIQUE RELATIVE À L'INCAPACITÉ DE TRAVAIL DE LONGUE DURÉE ET À L'INVALIDITÉ

Vu le règlement (UE) n°2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données) ;

Vu la loi du 3 décembre 2017 *portant création de l'Autorité de protection des données*, notamment l'article 114, modifié par la loi du 25 mai 2018 ;

Vu la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel*;

Vu la loi du 13 décembre 2006 *portant dispositions diverses en matière de santé*, en particulier l'article 42, § 2, 3°, modifié par la loi du 5 septembre 2018 ;

Vu la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, notamment les articles 5 et 15 ;

Vu la loi du 5 septembre 2018 *instituant le comité de sécurité de l'information et modifiant diverses lois concernant la mise en œuvre du Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE*, notamment l'article 97 ;

Vu la loi du 21 août 2008 *relative à l'institution et à l'organisation de la plate-forme eHealth* ;

Vu la demande d'autorisation de l'Ecole de santé publique de l'ULB ;

Vu le rapport d'auditorat de la Plate-forme eHealth du 31 mai 2022 ;

Vu le rapport de monsieur Bart Viaene ;

Émet, après délibération, la décision suivante, le 7 juin 2022 :

I. OBJET DE LA DEMANDE

1. La demande s'inscrit dans le cadre d'une recherche scientifique portant sur l'identification des facteurs santé pré-incapacité contribuant à expliquer l'incapacité de travail de longue durée ou l'invalidité. Cette étude est réalisée dans le cadre d'une thèse de doctorat au sein du Département d'Economie appliquée de l'Université Libre de Bruxelles (DULBEA) sous la supervision du Professeur Ilan Tojerow.
2. La réalisation de cette étude nécessite le couplage de données à caractère personnel relatives à la santé pseudonymisées communiquées par l'Agence intermutualiste et de données à caractère personnel pseudonymisées issues du datawarehouse marché du travail et protection sociale de la BCSS.
3. Les personnes concernées sont issues de la population qui a droit à une allocation d'incapacité de travail en Belgique, qui correspond à l'ensemble de la population belge en âge de travailler qui est inscrite dans l'une des mutuelles belges. L'objectif de cette étude est d'identifier les facteurs de santé pré-invalidité contribuant à expliquer l'entrée (et la durée du séjour) en incapacité de travail. Seul un petit pourcentage (environ 6%) de la population en âge de travailler entrera en incapacité de travail à un moment donné, les chercheurs doivent donc considérer un échantillon d'au moins 10% de toutes les personnes éligibles à l'incapacité de travail afin de pouvoir analyser la différence entre les personnes qui finissent par s'inscrire à ce programme et celles qui ne le font pas. Ce faisant, les chercheurs seraient en mesure d'identifier les facteurs de risque qui peuvent (partiellement) expliquer pourquoi une personne entrerait en incapacité de travail.
4. Les chercheurs ont besoin d'un sous-échantillon représentatif de 10 % extrait de l'ensemble de la population belge éligible à une incapacité de travail (16-64 ans) (5 % extrait en 2012 et 5 % extrait en 2013) et suivi entre 2006 et 2019. Sachant que la population belge en âge de travailler est d'environ 7,35 millions, il s'agit de recevoir un sous-échantillon d'environ 735.000 personnes.
5. La probabilité d'entrer en incapacité de travail est d'environ 6 %, il est donc nécessaire d'avoir suffisamment de variation parmi le pourcentage de personnes en incapacité dans l'échantillon. Dans la mesure où les demandeurs cherchent à identifier tous les facteurs de risque possibles (facteurs liés à la santé, pouvant être approximatifs par l'utilisation des prestations de soins et la consommation de médicaments) pouvant (partiellement) déterminer l'entrée en incapacité de travail, ils doivent s'assurer que les données couvrent la plupart des combinaisons possibles d'éléments. Cela signifie qu'ils doivent garantir toutes les combinaisons possibles de caractéristiques individuelles pour les personnes en invalidité (sexe, âge, région, niveau socio-économique, secteur d'activité, type de ménage, type et durée d'incapacité, pathologie, schémas de consommation de médicaments et de prestations de santé,...). Étant donné qu'ils analysent un échantillon relativement petit avec une grande variation de caractéristiques, il est nécessaire que le sous-échantillon soit suffisamment

représentatif de la population réelle, autrement les résultats ne seront pas suffisamment cohérents. De plus, comme la variable de résultat sera une variable dichotomique qui prend la valeur 1 si l'individu commence une incapacité de travail et 0 autrement pour chaque trimestre, travailler avec un échantillon insuffisamment grand impliquerait une forte probabilité d'avoir 0 entrée dans plusieurs cellules, cela pourrait invalider les résultats en présentant des modèles d'entrée irréalistes.

6. 10% est la taille minimale pour avoir une puissance statistique suffisante pour les estimations et la représentativité de la population. A partir de ce nombre, on peut se fier au Théorème Central du Limite, qui explique que si des individus sont choisis aléatoirement dans la population, la loi des grands nombres se chargera d'égaliser les moyennes d'un grand échantillon avec les moyennes de la population, assurant la représentativité de l'échantillon. En ce sens, plus l'échantillon est grand, mieux c'est, mais avec moins de 10 %, nous ne pouvons pas garantir une variabilité suffisante.
7. Suivant le principe de randomisation, avoir un sous-échantillon représentatif des 10 % de la population ciblée signifiera qu'environ 6 % de l'échantillon correspondra à des personnes en incapacité de travail (environ 44.100 individus). Les chercheurs ont besoin dans ce sous-échantillon d'une variabilité suffisante des caractéristiques. Par exemple, nous savons que parmi toutes les personnes en incapacité, un peu moins de 30 % entrent dans le programme en raison d'un problème de santé mentale. Ainsi, 30 % de ces 44.100, soit 13.230 individus de notre échantillon, auraient la santé mentale comme pathologie primaire pour entrer en incapacité de travail, et il en sera de même avec d'autres pathologies (la plupart représentant un pourcentage plus faible) et avec caractéristiques sociodémographiques. Les chercheurs doivent s'assurer que la variabilité des personnes en situation d'invalidité de l'échantillon correspond à la variabilité de la population, ils doivent travailler avec une taille d'échantillon qui ne permette à aucune caractéristique pertinente d'être sous-représentée.
8. Les résultats de cette étude mettront en lumière une préoccupation importante et croissante pour la Belgique ainsi que pour plusieurs pays de l'OCDE, et aideront à concevoir des politiques publiques qui pourraient avoir un impact social significatif. Pour cette raison, il est important d'obtenir les résultats les plus précis possibles. Plus l'échantillon est petit (et donc moins représentatif), plus le risque de ne pas rassembler toutes les informations pertinentes est élevé.

Procédure de sélection

9. La procédure proposée est d'extraire 5% de la population pour l'année 2012, 5% pour l'année 2013, de manière aléatoire, de la base de données BCSS, et de suivre tous ces individus entre 2006 et 2019, en utilisant des données trimestrielles. Les chercheurs proposent d'utiliser les années 2012 et 2013 mais il peut être envisagé d'utiliser d'autres années entre 2006 et 2019 si nécessaire. La BCSS est chargée d'assurer la représentativité de l'échantillon en l'extrayant de la population de manière aléatoire. Les chercheurs demandent à la BCSS d'extraire ce sous-échantillon et de l'envoyer à l'IMA.
10. Le schéma de couplage en annexe explique le trajet des données BCSS et IMA, les données pseudonymisées seront mises à disposition des chercheurs sur le serveur de l'IMA.

II. COMPÉTENCE

a. Données issues du Datawarehouse Marché du travail et protection sociale

11. En vertu de l'article 5, § 1er, de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, la Banque Carrefour de la sécurité sociale recueille des données à caractère personnel auprès des institutions de sécurité sociale, les enregistre, procède à leur agrégation et les communique aux personnes qui en ont besoin pour la réalisation de recherches pouvant être utiles à la connaissance, à la conception et à la gestion de la protection sociale. Il s'agit en l'occurrence d'une communication de données à caractère personnel qui, en vertu de l'article 15, § 1er, de la loi du 15 janvier 1990, doit faire l'objet d'une délibération de la chambre sécurité sociale et santé du Comité de sécurité de l'information.
12. La chambre sécurité sociale et santé du Comité de sécurité de l'information s'estime dès lors compétente pour se prononcer sur la présente demande.

III. EXAMEN

A. ADMISSIBILITÉ

a. Données issues du Datawarehouse Marché du travail et protection sociale

13. Selon l'article 6 du RGPD, le traitement de données à caractère personnel n'est licite que si, et dans la mesure où, au moins une des conditions mentionnées est remplie.
14. Le traitement précité est licite en ce qu'il est nécessaire au respect d'une obligation légale à laquelle est soumise le responsable du traitement, conformément à l'article 6, 1, c), du RGPD, à savoir l'article 2 du décret du 7 novembre 2013 définissant le paysage de l'enseignement supérieur et l'organisation académique des études. Cet article prévoit que les établissements de l'enseignement supérieur en Communauté française ont, en plus de la mission d'octroyer les titres et grades académiques sanctionnant les études supérieures et de délivrer les diplômes et certificats correspondants, trois missions à remplir. Une de ces trois missions complémentaires est de participer à des activités individuelles ou collectives de recherche, d'innovation ou de création et assurer ainsi le développement, la conservation et la transmission des savoirs et du patrimoine culturel, artistiques et scientifique.
15. A la lumière de ce qui précède, le comité de sécurité de l'information est par conséquent d'avis qu'il existe un fondement admissible pour le traitement des données à caractère personnel pseudonymisées relatives à la santé envisagé.

B. PRINCIPES RELATIFS AU TRAITEMENT DES DONNÉES À CARACTÈRE PERSONNEL

1. FINALITÉ

16. Selon l'article 5 du RGPD, les données à caractère personnel doivent être traitées de manière licite, loyale et transparente au regard de la personne concernée. Elles doivent être collectées pour des finalités déterminées, explicites et légitimes, et ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités.
17. Les données utilisées dans le cadre de cette étude, le sont dans le cadre d'un traitement ultérieur par rapport aux finalités pour lesquelles elles étaient initialement collectées.
18. L'objectif de l'étude est de croiser les données relatives au marché du travail et aux autres branches de la sécurité sociale avec les données de remboursement des prestations de soins de santé pour identifier les facteurs santé pré-incapacité contribuant à expliquer l'incapacité de travail de longue durée ou à l'invalidité.

a. Données issues du datawarehouse marché du travail et protection sociale

19. En vertu de l'article 5 de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, la BCSS recueille des données sociales auprès des institutions de sécurité sociale, les enregistre, procède à leur agrégation et les communique à des personnes qui en ont besoin pour la réalisation de recherches pouvant être utiles à la connaissance, à la conception et à la gestion de la protection sociale.
20. En vertu de l'article 5, §1^{er}, b) du RGPD, le traitement ultérieur à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques n'est pas considéré, conformément à l'article 89, paragraphe 1, comme incompatible avec les finalités initiales.
21. Cette étude est réalisée dans le cadre d'une recherche scientifique effectuée au sein de l'Université libre de Bruxelles.

b. Données communiquées par l'Agence intermutualiste

22. Les unions nationales des Organismes Assureurs (OA) disposent de données relatives à leurs membres, dans le cadre de l'exécution de leur mission légale liée à l'AMI. L'objectif de l'Agence Intermutualiste est d'analyser les données collectées dans le cadre de leurs missions et de fournir l'information en la matière (Art. 278 alinéa 1 de la Loi Programme 24.2/2002). Par ailleurs, cette étude scientifique peut contribuer aux missions légales des OA tel que défini par la loi du 6 août 1990 (art. 3a et b), à savoir la participation à la mise en œuvre de l'assurance obligatoire soins de santé et indemnités, régie par la loi coordonnée du 14 juillet 1994.
23. L'article 5, §1^{er} b) du RGPD prévoit que le traitement ultérieur à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques n'est pas considéré, conformément à l'article 89, paragraphe 1, comme incompatible avec les finalités initiales (limitation des finalités).

24. Au vu des objectifs du traitement tels que décrits ci-dessus, le Comité de sécurité de l'information considère que le traitement des données à caractère personnel envisagé poursuit bien des finalités déterminées, explicites et légitimes.

2. PROPORTIONNALITÉ

25. L'article 5, §1er du RGPD dispose que les données à caractère personnel doivent être adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données).
26. Les chercheurs ont besoin d'un sous-échantillon représentatif de 10 % extrait de l'ensemble de la population belge éligible à une incapacité de travail (16-64 ans) (5 % extrait en 2012 et 5 % extrait en 2013) et suivi entre 2006 et 2019. Sachant que la population belge en âge de travailler est d'environ 7,35 millions, il s'agit de recevoir un sous-échantillon d'environ 735.000 personnes. Un sous-échantillon représentatif des 10 % de la population ciblée signifiera qu'environ 6 % de l'échantillon correspondra à des personnes en incapacité de travail (environ 44.100 individus).
27. Le demandeur déclare que le traitement des différentes données à caractère personnel pseudonymisées demandées est nécessaire pour les raisons suivantes :

a. Données issues du Datawarehouse Marché du travail et protection sociale

28. Les variables relatives aux **caractéristiques personnelles et socio-économiques** des personnes concernées pour la période 2006-2019 (par trimestre) :
- le NISS pseudonymisé : nécessaire pour regrouper les informations relatives à un même individu ;
 - la date de naissance (mois/année) : variable à caractère personnel élémentaire comme variable de contrôle dans les estimations. Les chercheurs n'ont pas besoin de la date de naissance exacte de l'individu, mais ils ont besoin de l'âge au cours du trimestre. Il est important d'avoir l'âge comme variable continue car l'apparition de maladies pouvant conduire une personne à l'invalidité est étroitement et positivement corrélée avec l'âge. Il est également important de prendre en compte la distance qui sépare chaque individu de l'âge de pension ;
 - le sexe : variable à caractère personnel élémentaire comme variable de contrôle dans les estimations. Le sexe est largement reconnu comme un facteur d'impact déterminant dans de nombreux phénomènes sociaux. Dans le cas de la santé et des incapacités au travail, plusieurs études démontrent ces différences ;
 - catégories d'études : variables permettant de différencier les effets en fonction du niveau et du type d'étude du demandeur d'emploi ;
 - première nationalité (en classes) et nationalité (en classes) : variable à caractère personnel élémentaire comme variable de contrôle dans les estimations. Le phénomène de l'immigration devient de plus en plus fréquent, notamment en Belgique. C'est pourquoi cette variable permet de savoir s'il existe des différences entre les immigrés et les

nationaux, puisqu'ils sont également confrontés à des réalités socio-économiques et d'intégration différentes qui peuvent avoir un impact sur la santé ;

- le code d'arrondissement de la commune du domicile : variable à caractère personnel élémentaire comme variable de contrôle dans les estimations. Il est important de prendre en compte les différences socio-économiques et géographiques dérivées de la commune de résidence. Le type de marché du travail (étendue et disponibilité), ainsi que l'offre de ressources en santé dans chaque lieu, peuvent entraîner des différences très importantes qui affectent notre domaine d'études ;
- le statut social (en classes) : variable à caractère personnel élémentaire comme variable de contrôle dans les estimations. Variables nécessaires pour prendre en compte d'éventuels effets de revenus ;
- Indicateur statistique :.....
- Régime (secteur privé, secteur public, enseignement) : variable à caractère personnel élémentaire comme variable de contrôle dans les estimations. Indique si la personne qui bénéficie de l'interruption de carrière est employée dans le secteur public, le privé ou l'enseignement.
- le code de composition du ménage : variable de base nécessaire dans toute analyse sur l'incapacité de travail (afin de distinguer entre différents types de prestations selon le type d'individu).
- le code de la position dans le ménage suivant la typologie LIPRO
- revenu annuel (en classes)
- date de décès (mois/année) : utile pour connaître la gravité de la maladie.
- le code de nomenclature de la position socio-économique : variables permettant d'identifier précisément le statut de chaque individu et d'observer les changements de statuts des individus entre différentes périodes ;
- indication si la personne était dans une situation d'emploi au cours du trimestre (oui/non) ;
- salaire (en classes) : le montant de la rémunération " ordinaire " du trimestre, c'est-à-dire toute rémunération brute (sauf les charges fiscales) qui ne constitue pas une indemnité de rupture, une prime ou un salaire d'attente (voir autres codes salariaux).
- indépendant (oui/non) ;
- revenu annuel de l'INASTI (base annuelle – en classes) ;
- chômage (oui/non) : variable indicatrice si la personne a connu une période de chômage pendant le trimestre si la personne perçoit des allocations de chômage de l'ONEm ;
- activation ONEm (oui/non) : position sur le marché du travail d'actif occupé en combinaison avec une activation de l'ONEm. On peut donc savoir si la personne perçoit des allocations de chômage de l'ONEm ;
- Incapacité (oui/non) : variable indicatrice si la personne a connu une période d'incapacité pendant le trimestre ;

29. Les variables relatives à l'incapacité de travail des personnes concernées pour la période 2006-2019 (par trimestre) :

- date de début d'invalidité (mois/années) : cette variable indique la date de début en invalidité. Les chercheurs n'ont pas besoin de la date exacte, seulement du trimestre de l'année. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation ;

- date de début de la période de paiement (mois/année) : cette variable indique la date de début du paiement, qu'elle puisse être différente du début en incapacité de travail, et il est important d'analyser cette différence, si elle existe ;
- date de séance (mois/année) ;
- date de début de la maladie, date de début de l'incapacité primaire (mois/année) : cette variable indique la date de début de la maladie, qu'elle puisse être différente du début en incapacité de travail, et il est important d'analyser cette différence, si elle existe ;
- le code de la décision du conseil médical de l'invalidité : l'invalidité est reconnue par le Conseil médical de l'invalidité sur la base d'un rapport détaillé établi par le médecin-conseil de l'organisme assureur. Le Conseil doit prendre une décision dans les quatre dernières semaines de la période d'incapacité de travail primaire ;
- les jours indemnisés : variable nécessaire pour connaître l'importance/ la durée de l'invalidité et ainsi, distinguer entre différents types d'invalidité (courte durée >< longue durée, occurrences multiples, etc.) ;
- Montant indemnité : variable nécessaire pour connaître l'importance/ la durée de l'invalidité et ainsi, distinguer entre différents types d'invalidité ;
- R exclus : variable permettant d'identifier le statut exact de l'individu à la fin du trimestre ;
- date de début et de fin d'incapacité de travail (mois/année) ;
- montant de l'incapacité de travail primaire (en classes) ;
- la nature de l'indemnité ;
- le nombre de jours d'incapacité de travail : pour connaître l'importance/la durée de l'incapacité ;
- le type d'indemnité et le type de jours : variables permettant de différencier les effets de différents types de bénéficiaires de l'allocation d'incapacité ;
- le code indemnité, le code médical, le code sortie du régime : variables permettant de différencier les effets de différents types de bénéficiaires de l'allocation d'invalidité, ainsi de différentes maladies et divers motifs de sortie du statut.

30. Les variables relatives à **l'emploi** des personnes concernées pour la période 2006-2019 (par trimestre) :

- date de début (mois/année) : date de début de l'emploi dans le cadre d'une mesure d'activation. Variable nécessaire pour identifier le régime auquel chaque individu est soumis en fonction des évolutions législatives ;
- contrat de travail : permet d'identifier le statut des individus ayant souscrit à un contrat de d'interruption de carrière ou de crédit-temps ;
- classe de travailleur (ouvrier/employé/fonctionnaire/autre), classe de travailleur spécial (le code), classe salariale, classe temps partiel (en pourcentage) : variables permettant d'obtenir des résultats se basant sur le parcours professionnel détaillé de l'individu au cours d'une période donnée. Ces variables permettent également de distinguer les effets de différentes situations professionnelles actuelles ou passées de l'individu. Certaines variables (notamment la commission paritaire applicable) permet d'identifier le respect des conditions pour bénéficier de certaines allocations (notamment en ce qui concerne les prépensions). D'autres variables (telles que la classe salariale ou le salaire journalier) permettent de prendre en compte d'éventuels effets de revenus ;

- le code de la commission paritaire : indique le numéro de la commission paritaire, qui est déterminé en fonction de l'activité économique de l'entreprise. Ce qui pourrait sans doute être très important dans l'évolution de l'incapacité de travail ;
- le nombre total de jours assimilés du trimestre : cette variable fait référence au cheminement de travail de chaque individu, ce qui pourrait également avoir un effet à la fois sur la détérioration de la santé de la personne et sur la décision d'amorcer une démarche pour demander une invalidité ;
- le pourcentage de temps partiel : le type de journée de travail et le nombre d'heures travaillées pourrait avoir un effet à la fois sur la détérioration de la santé de la personne (burnout, problèmes de santé mentale et musculo-squelettiques) et sur la décision d'amorcer une démarche pour demander une invalidité ;
- le salaire journalier (en classes) : le salaire reflète le pouvoir d'achat de l'individu, information nécessaire à contrôler dans les études économétriques à caractère socio-économique, puisqu'il peut générer des disparités importantes ;
- code qui indique si le travailleur a été occupé via le régime des titres-services : il a été démontré que ces travailleurs sont particulièrement susceptibles d'avoir une invalidité et qu'ils constituent donc un sous-groupe important à étudier lorsqu'on essaie de prédire les modèles d'entrées en incapacité de travail ;
- le code nace (3 premiers chiffres) : indique le secteur dans lequel une personne était employée, ce qui pourrait sans doute être très important dans la conduite d'une incapacité de travail ;
- équivalent temps plein, journées assimilées incluses : l'équivalent temps plein donne une indication du degré d'occupation d'un emploi au cours du trimestre. Ce qui pourrait sans doute être très important dans l'évolution de l'incapacité de travail.

b. Données pseudonymisées communiquées par l'Agence intermutualiste

- 31.** Les variables communes des personnes concernées pour la période 2006-2019 (par trimestre) :
- le numéro d'identification du bénéficiaire (soins de santé, pharmanet, hospitalisation) : variables indispensables pour regrouper les informations s'apparentant à un même individu ;
 - variable dérivée de la date début de prestation et de la date début de délivrance du médicament (trimestre et année) ;
 - le code de nomenclature (soins de santé, pharmanet) : le code nomenclature permet de déterminer pour quelle prestation de soins un patient a été remboursé entièrement ou partiellement par l'assurance-maladie ;
 - le montant du remboursement (soins de santé, pharmanet) par tranches en euros.
- 32.** Les variables liées à la consommation des médicaments des personnes concernées pour la période 2006-2019 (par trimestre)
- variable dérivée code nomenclature (soins de santé, pharmanet) : cette variable indique si la prestation est un médicament ;
 - pharmacie (soins de santé, pharmanet) : pharmacie publique ou hospitalière ;
 - ATC code niveau 2 (soins de santé, pharmanet) : le système de classification ATC (anatomique, thérapeutique et chimique) est un classement international des médicaments. Élaboré par l'OMS, il attribue à chaque principe actif un code dans un

classement hiérarchique selon une répartition anatomique, thérapeutique et chimique. Le premier niveau indique le groupe anatomique principal. Le deuxième niveau indique le sous-groupe thérapeutique. Le code ATC est nécessaire afin de savoir quel médicament a été délivré en pharmacie hospitalière. Les chercheurs ont besoin de classer les médicaments selon leur code ATC. Dans la plupart des cas, le premier et le deuxième niveau du code ATC est suffisant. Cela nous permet de se faire une idée du type de médicament consommé et du type de maladie, sans savoir quel est le médicament exact. Dans le cas des médicaments délivrés, les chercheurs veulent avoir les informations sur la quantité consommée en DDD (Daily Defined Dose) au niveau ATC car c'est un moyen simple et d'usage international. Étant donné que le caractère de cette étude n'est pas médical/pharmacologique, les chercheurs n'ont pas besoin de connaître la dose exacte de médicament qu'un individu consomme, mais plutôt une mesure standard qui nous permet de faire des comparaisons à un niveau agrégé. Le nombre de DDD peut être calculé comme suit : $DDD = SS00050 * DPP$;

- DDD (Daily defined dose) : dans le cas des médicaments délivrés, nous voulons avoir les informations sur la quantité consommée en DDD (Daily Defined Dose) au niveau ATC car c'est un moyen simple et d'usage international. Étant donné que l'étude n'est pas médicale/pharmacologique, les chercheurs n'ont pas besoin de connaître la dose exacte de médicament qu'un individu consomme, mais plutôt une mesure standard qui nous permet de faire des comparaisons à un niveau agrégé. Le nombre de DDD peut être calculé comme suit : $DDD = SS00050 * DPP$.

33. Les variables liées aux visites chez le médecin des personnes concernées pour la période 2006-2019 (par trimestre) :

- Prestation B (visite chez un médecin) ;
- Généraliste/ spécialiste : cette variable indique le code profession du prestataire de soins qui a effectué la prestation. Ce code à deux chiffres indique quel type de prestataire a effectué la prestation de soins. Cette variable indique si la présentation est une visite chez le médecin généraliste ou avec un médecin spécialiste. Les chercheurs doivent connaître le type de médecin que le patient consulte, en vue d'identifier des effets hétérogènes selon le type de maladie dont souffre l'individu ;
- Spécialiste (par catégorie) : les chercheurs n'ont pas besoin d'informations sur toutes les spécialisations existantes. Dans le cas où la consultation se fait avec un médecin spécialiste, les chercheurs souhaitent savoir s'il appartient à l'une des 8 (+2) spécialisations. Ces 8 (+2) spécialités puisqu'elles sont les plus courantes avant d'entrer en incapacité de travail de longue durée. Il est également nécessaire de connaître le nombre de fois où un individu a consulté un kinésithérapeute, au cours d'un trimestre donné. Un kinésithérapeute n'est pas considéré comme un médecin spécialiste, il est important de prendre également en compte la consommation de cette ressource, puisqu'elle peut être associée à la présence des maladies musculo-squelettiques, l'une des plus fréquentes chez les personnes en invalidité. Les chercheurs ont besoin des informations sur le nombre de visites chez un psychologue. Comme dans le cas du kinésithérapeute, un psychologue n'est pas un médecin spécialiste, mais il est également intéressant d'avoir cette information, car une consultation avec un psychologue pourrait être liée à des problèmes de santé mentale, qui sont également très fréquents chez les personnes en incapacité de travail.

- 34.** Les variables liées aux **hospitalisations** des personnes concernées pour la période 2006-2019 (par trimestre) :
- prestation C (hospitalisation) : cette variable indique si la présentation est une hospitalisation ;
 - date d'admission de l'hospitalisation (mois/année) : début du séjour hospitalier, 1er jour d'hospitalisation. La date d'admission mentionnée dans la facturation d'une prestation hospitalière ;
 - admission pour une hospitalisation de jour (mois/année) : la variable "date d'admission" est rarement remplie. Il faut alors employer cette variable ;
 - date de sortie de l'hôpital (mois/année) : la date de sortie de l'hôpital est celle reprise lors de la facturation d'une prestation hospitalière si elle est connue ;
 - nombre de jours d'hospitalisation : longueur du séjour (en jours) d'hospitalisation dans le trimestre de référence.
- 35.** Le principe de proportionnalité implique que le traitement doit en principe être réalisé au moyen de données anonymes. Cependant, si la finalité ne peut être réalisée au moyen de données anonymes, des données à caractère personnel pseudonymisées peuvent être traitées. Vu la nécessité de réaliser des analyses très détaillées à l'aide de ces données, le demandeur a besoin d'avoir accès à des données pseudonymisées afin d'être en mesure de réaliser des analyses qu'il ne serait pas possible de réaliser à l'aide de données anonymes. Cette finalité justifie donc le traitement de données à caractère personnel pseudonymisées.
- 36.** Le Comité constate qu'une *small cell risk analysis* sera réalisée par l'Agence intermutualiste dès qu'un accord aura été trouvé avec l'ULB.
- 37.** Conformément à l'article 5, §1er, e), les données à caractère personnel doivent être conservées sous une forme permettant l'identification des personnes concernées pendant une durée n'excédant pas celle nécessaire au regard des finalités pour lesquelles elles sont traitées.
- 38.** Les données issues du couplage des bases de données seront conservées sur un serveur sécurisé de l'ULB jusqu'au 30 juin 2025. Les données seront ensuite détruites. Cette durée est nécessaire pour analyser les données, faire des estimations et répondre aux questions possibles pour publier les résultats de cette étude dans une revue scientifique. De plus, il correspond à la même période de temps pour laquelle l'INAMI a accordé son financement pour réaliser ce projet.
- 39.** Les données sont tenues à la disposition des chercheurs sur l'infrastructure de l'AIM, gérée par la TTP OA (BCSS) jusqu'à la fin du projet. Une fois les projets de recherche terminés et le rapportage des résultats effectué, les données disponibles sont supprimées. La gestion et la suppression des données est assurée par les collaborateurs de l'AIM et de la TTP OA (BCSS), et ce selon les exigences explicites mentionnées dans l'autorisation.
- 40.** Le Comité de sécurité de l'information estime que ce délai de conservation est raisonnable.

3. TRANSPARENCE

41. Conformément à l'article 12 du RGPD, le responsable du traitement doit prendre des mesures appropriées pour fournir toute information en ce qui concerne le traitement à la personne concernée d'une façon concise, transparente, compréhensible et aisément accessible, en des termes clairs et simples. Les informations sont fournies par écrit ou par d'autres moyens y compris, lorsque c'est approprié, par voie électronique.
42. En vertu de l'article 14, al. 5, b) du RGPD, le responsable du traitement est dispensé de cette obligation lorsque la communication de telles informations se révèle impossible ou exigerait des efforts disproportionnés, en particulier pour le traitement à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques sous réserve des conditions et garanties visées à l'article 89, paragraphe 1, du RGPD.
43. Les chercheurs ne sont pas en mesure de contacter toutes les personnes concernées étant donné qu'il s'agit d'un traitement ultérieur de données pseudonymisées ne comportant pas de données de contact.
44. Le Comité de sécurité de l'information est d'avis qu'il existe suffisamment de transparence quant au traitement envisagé.

4. MESURES DE SÉCURITÉ

45. Selon l'article 5, §1er, f) du RGPD, les données à caractère personnel doivent être traitées de façon à garantir une sécurité appropriée des données à caractère personnel, y compris la protection contre le traitement non autorisé ou illicite et contre la perte, la destruction ou les dégâts d'origine accidentelle, à l'aide de mesures techniques ou organisationnelles appropriées (intégrité et confidentialité).
46. Pour garantir la confidentialité et la sécurité du traitement de données, tout organisme qui conserve, traite ou communique des données à caractère personnel est tenu de prendre des mesures dans les onze domaines d'action suivants liés à la sécurité de l'information: politique de sécurité; désignation d'un conseiller en sécurité de l'information; organisation et aspects humains de la sécurité (engagement de confidentialité du personnel, information et formations régulières du personnel sur le thème de la protection de la vie privée et sur les règles de sécurité); sécurité physique et de l'environnement; sécurisation des réseaux; sécurisation logique des accès et des réseaux; journalisation, traçage et analyse des accès; surveillance, revue et maintenance; système de gestion des incidents de sécurité et de la continuité (systèmes de tolérance de panne, de sauvegarde, ...); documentation
47. Le Comité constate que le schéma de couplage de données prévoit l'intervention du TTP eHealth et du TTP OA BCSS.
48. Le Comité rappelle que le rôle de tiers de confiance de la plateforme eHealth est défini par l'article 5, 8° de la loi du 21 août 2008 *relative à l'institution et à l'organisation de la plateforme eHealth* : « en tant qu'organisme intermédiaire au sens d'une organisation autre que le responsable du traitement de données à caractère personnel non pseudonymisées, qui est chargée de leur pseudonymisation, recueillir, agréger, coder ou anonymiser et mettre à disposition des données utiles à la connaissance, à la conception, à la gestion et à la prestation

de soins de santé; la plate-forme eHealth ne pourra conserver les données à caractère personnel traitées dans le cadre de cette mission que pour la durée nécessaire à leur codification ou anonymisation; la plate-forme eHealth peut cependant conserver le lien entre le numéro d'identification réel d'une personne concernée et le numéro d'identification codé qui lui a été attribué, si le destinataire des données à caractère personnel codées en fait la demande d'une façon motivée, moyennant une autorisation [de la chambre sécurité sociale et santé du comité de sécurité de l'information; la plate-forme eHealth peut uniquement réaliser cette mission à la demande d'une chambre législative, d'une institution de sécurité sociale, de la fondation visée à l'article 45quinquies de l'arrêté royal n° 78 du 10 novembre 1967 relatif à l'exercice des professions de soins de santé, de l'Agence intermutualiste, du Centre fédéral d'expertise des soins de santé, d'un ministre fédéral, d'un service public fédéral ou d'une institution publique dotée de la personnalité juridique qui relève des autorités fédérales; le Roi peut, par arrêté délibéré en Conseil des Ministres et après avis de la Commission de la protection de la vie privée et du Comité de gestion, élargir la liste des instances qui peuvent faire appel à la plate-forme eHealth comme organisation intermédiaire.

49. Le Comité constate que les données pseudonymisées seront mises à disposition des chercheurs via un serveur sécurisé de l'Agence Intermutualiste. Le Comité rappelle que l'IMA et l'ULB (DULBEA) sont tenus de signer une convention décrivant la mise à disposition de ces données, les mesures de sécurité et les responsabilités de chaque intervenant dans le cadre de cette étude.
50. Le Comité rappelle que les données issues du couplage (les résultats des recherches) devront être conservées sur un serveur sécurisé de l'ULB et uniquement accessible aux chercheurs participants à l'étude.
51. Le Comité constate que les données seront traitées sous la responsabilité d'un professionnel des soins de santé.
52. Le Comité constate que les chercheurs de l'ULB s'engagent à respecter la confidentialité de la recherche ainsi qu'à respecter des obligations strictes de confidentialité en matière de données à caractère personnel traitées dans le cadre de leurs prestations.
53. Le Comité constate que les chercheurs déclarent qu'une analyse d'impact relative à la protection des données sera réalisée avant la mise à disposition des données. Le Comité demande que les résultats de cette analyse lui soient communiqués dans les meilleurs délais.
54. La chambre sécurité sociale et santé rappelle qu'en vertu de l'article 9 de la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel*, le responsable du traitement prend les mesures supplémentaires suivantes lors du traitement de données génétiques, biométriques ou des données concernant la santé :
 - 1° les catégories de personnes ayant accès aux données à caractère personnel, sont désignées par le responsable du traitement ou, le cas échéant, par le sous-traitant, avec une description précise de leur fonction par rapport au traitement des données visées;

2° la liste des catégories des personnes ainsi désignées est tenue à la disposition de l'autorité de contrôle compétente par le responsable du traitement ou, le cas échéant, par le sous-traitant;

3° il veille à ce que les personnes désignées soient tenues, par une obligation légale ou statutaire, ou par une disposition contractuelle équivalente, au respect du caractère confidentiel des données visées.

Par ces motifs,

la chambre sécurité sociale et santé du comité de sécurité de l'information

sous réserve de la réalisation d'une analyse d'impact relative à la protection des données avant la mise à disposition des données aux chercheurs,

conclut que

la communication des données à caractère personnel telle que décrite dans la présente délibération est autorisée moyennant le respect des mesures de protection de la vie privée qui ont été définies, en particulier les mesures en matière de limitation de la finalité, de minimisation des données, de limitation de la durée de conservation des données et de sécurité de l'information.

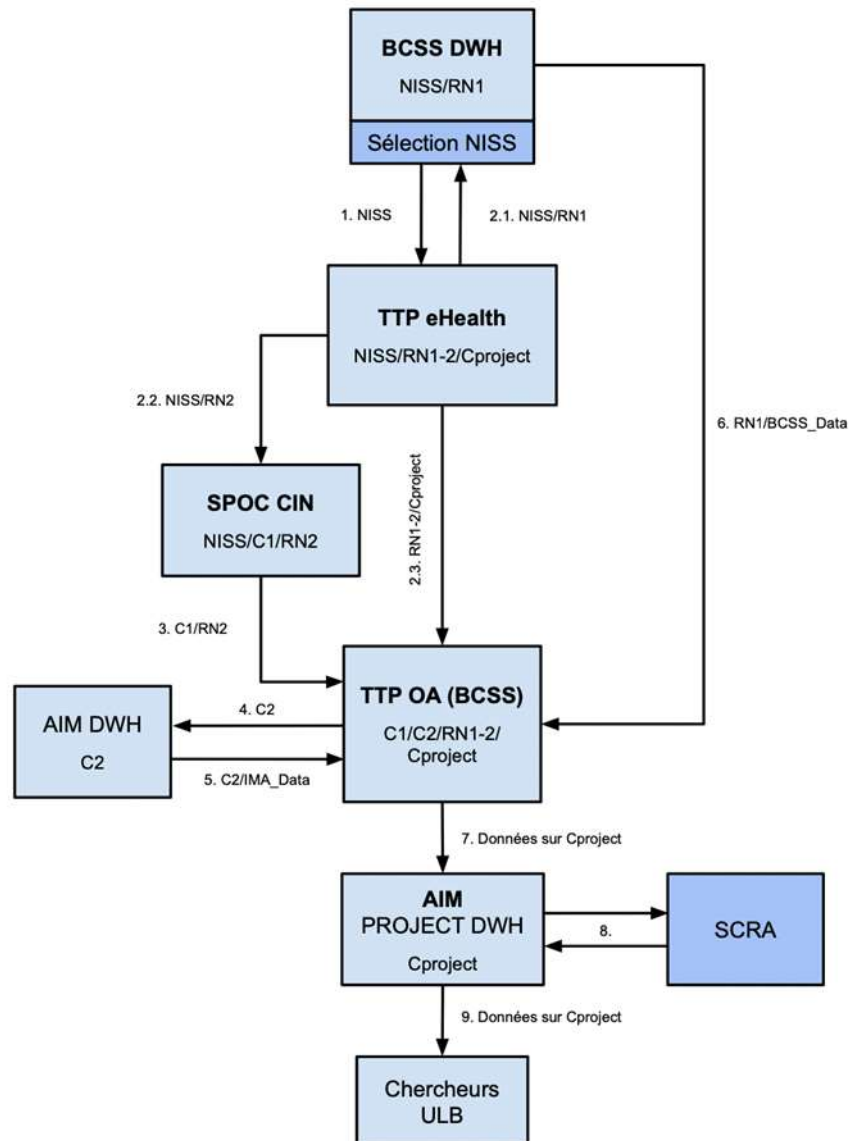
Bart VIAENE
Président

Le siège de la chambre sécurité sociale et santé du comité de sécurité de l'information est établi dans les bureaux de la Banque Carrefour de la sécurité sociale, à l'adresse suivante: Quai de Willebroeck 38 - 1000 Bruxelles (tél. 32-2-741 83 11).

Annexe 1

Schéma de couplage

Identification et analyse des facteurs santé
pré-incapacité: flux des données BCSS - AIM
(P_ARBO01)



1. Dans le cadre de cette étude, une sélection de personnes est faite par la BCSS, et la liste de NISS de ces personnes est envoyée au TTP eHealth.
2. Le TTP eHealth:
 - 2.1. Produit à partir de la liste de NISS une première série de codes RN et envoie la liste de conversion NISS/RN1 à la BCSS.
 - 2.2. Produit à partir de la liste de NISS une deuxième série de codes RN et envoie la liste de conversion NISS/RN2 au SPOC CIN.
 - 2.3. Produit à partir de la liste de NISS les codes Cproject et envoie la liste de conversion NISS/RN1/RN2/Cproject à la TTP OA (BCSS).
3. Le SPOC NIC convertit chaque NISS de la liste NISS/RN2 en C1 et envoie la liste C1/RN2 à la TTP OA (BCSS).
4. La TTP OA (BCSS) convertit les C1 en C2 et met à disposition les C2 de l'AIM en DWH.
5. L'AIM effectue la sélection des données AIM et transfère les données (C2/IMA_Data) à la TTP OA (BCSS).
6. La BCSS effectue la sélection des données BCSS et transfère les données sur RN (RN1/BCSS_Data) à la TTP OA (BCSS).
7. La TTP-OA (BCSS) remplace dans les données AIM le C2 par le Cprojet spécifique et code les données reçues de la part de la BCSS également sur ce Cprojet. Toutes les données sont placées sur Cprojet dans l'espace Projet protégé du DWH AIM.
8. Small Cells Risk Analysis à exécuter par le KCE.
9. Les données sont mises à disposition des chercheurs ULB sur Cproject.

Annexe 2
Liste des données pseudonymisées

Données pseudonymisées communiquées par l'AIM/IMA					
Nom de la variable	Code de la variable	Banque de données	Justification	Date et périodicité de les données	Valeurs possibles
Variables communes					
Identification bénéficiaire	SS00010	Soins de santé	Variable indispensable afin de regrouper les informations s'apparentant à un même individu.	Trimestriel 2006-2019	Numéro
	SS00010	Pharmanet			
	SS00010	Hospitalisation			
Trimestre et année	Variable dérivée de « date début de prestation » = SS00015	Soins de santé	Cette variable indique la date de prestation de soins de santé. Nous n'avons pas besoin de la date exacte de la prestation, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.	Trimestriel 2006-2019	T/AAAA
	Variable dérivée de « date début délivrance » = SS00015	Pharmanet	Cette variable indique le date de délivrance du médicaments. Nous n'avons pas besoin de la date exacte de la prestation, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.	Trimestriel 2006-2019	T/AAAA
Nombre de prestation par trimestre	Variable dérivée de « code nomenclature » = SS00020	Soins de santé	Le code nomenclature permet de déterminer pour quelle prestation de soins un patient a été remboursé entièrement ou partiellement par l'assurance-maladie.	Trimestriel 2006-2019	Numéro (1,2,3,4,...)
	Variable dérivée	Pharmanet			
Montant de remboursement	SS00060	Soins de santé	Cette variable indique le montant remboursé à un patient par l'organisme assureur.	Trimestriel 2006-2019	Par tranches en euros
	SS00060	Pharmanet			
Variables liées à la consommation de médicaments					
Prestation A (médicament)	Variable dérivée de « code	Soins de santé	Cette variable indique si la présentation est un médicament.	Trimestriel 2006-2019	1 : la prestation est a medecin

	nomenclature » = SS00020				0 : autrement
	Variable dérivée	Pharmanet			
Pharmacie	Variable dérivée	Soins de santé	Cette variable indique si la présentation est un médicament délivré en pharmacie publique ou en pharmacie hospitalière. En pharmacie publique si la variable vient de « Pharmanet » et en pharmacie hospitalière si la variable vient de « soins de santé ».	Trimestriel 2006-2019	1 : délivré en pharmacie hospitalière 2 : délivré en pharmacie publique
		Pharmanet			
ATC code niveau 2	ATC_ANAT_L - Code ATC niveau 2	Soins de santé (les médicaments délivrés en pharmacie hospitalière)	<p>Cette variable indique le code ATC. Le système de classification ATC (anatomique, thérapeutique et chimique) est un classement international des médicaments. Élaboré par le WHO, il attribue à chaque principe actif un code dans un classement hiérarchique selon une répartition anatomique, thérapeutique et chimique. Le premier niveau indique le groupe anatomique principal. Le deuxième niveau indique le sous-groupe thérapeutique.</p> <p>Le code ATC est nécessaire afin de savoir quel médicament a été délivré en pharmacie hospitalière. Pour notre étude, nous avons besoin de classer les médicaments selon leur code ATC. Dans la plupart des cas, le premier et le deuxième niveau du code ATC est suffisant. Cela nous permet de nous faire une idée du type de médicament consommé et du type de maladie, sans savoir quel est le médicament exact.</p>	Trimestriel 2006-2019	A01 /.../ V20
		Pharmanet (les médicaments délivrés en pharmacie publique)			
DDD (Daily defined dose)	Variable dérivée de « quantité » = SS00050	Soins de santé (les médicaments délivrés en pharmacie hospitalière)	Dans le cas des médicaments délivrés, nous voulons avoir les informations sur la quantité consommée en DDD (Daily Defined Dose) au niveau ATC car c'est un moyen simple et d'usage international. Étant donné que le caractère de notre étude n'est pas médical/pharmacologique, nous n'avons pas besoin de	Trimestriel 2006-2019	numéro
		Pharmanet (les médicaments délivrés en pharmacie publique)			

			<p>connaître la dose exacte de médicament qu'un individu consomme, mais plutôt une mesure standard qui nous permet de faire des comparaisons à un niveau agrégé.</p> <p>Le nombre de DDD peut être calculé comme suit : $DDD = SS00050 * DPP$.</p>		
Variables liées aux visites chez le médecin					
Prestation B (visité chez un médecin)	Variable dérivée	Soins de santé	Cette variable indique si la présentation est une visite chez le médecin.	Trimestriel 2006-2019	1 : la prestation est une visite et consultation chez le médecin 0 : autrement
Généraliste/ spécialiste	Variable dérivée de « PRACTITIONER _CAT-Code profession prestataire de soins »	Soins de santé	<p>Cette variable indique le code profession du prestataire de soins qui a effectué la prestation. Ce code à deux chiffres indique quel type de prestataire a effectué la prestation de soins.</p> <p>Cette variable indique si la présentation est une visite chez le médecin généraliste ou avec un médecin spécialiste. Nous devons connaître le type de médecin que le patient consulte, cela peut nous aider à identifier des effets hétérogènes selon le type de maladie dont souffre l'individu.</p>	Trimestriel 2006-2019	1 : médecin généraliste 2 : médecin spécialiste
Spécialiste	Variable dérivée de « PRACTITIONER _CAT-Code profession prestataire de soins »	Soins de santé	Nous n'avons pas besoin d'informations sur toutes les spécialisations existantes. Dans le cas où la consultation se fait avec un médecin spécialiste, nous souhaitons savoir si vous appartenez à l'une des 8 (+2) spécialisations. Nous nous intéressons particulièrement à ces 8 (+2) spécialités puisqu'elles sont les plus courantes avant d'entrer en incapacité de travail de longue durée. Nous voulons aussi le nombre de fois qu'un individu a consulté un kinésithérapeute, au cours d'un trimestre donné. Un	Trimestriel 2006-2019	0: médecin généraliste 1: orthopédie 2: psychiatrique 3: physiques médecin & rehab 4: anesthésiste 5: neurologue 6: cardiologie 7: médecin interniste 8: chirurgienne 9: kinés 10: psychologue 11: autre

			kinésithérapeute n'est pas considéré comme un médecin spécialiste, mais nous considérons qu'il est important de prendre également en compte la consommation de cette ressource, puisqu'elle peut être associée à la présence des maladies musculo-squelettiques, l'une des plus fréquentes chez les personnes en invalidité. Nous voulons également avoir des informations sur le nombre de visites chez un psychologue. Comme dans le cas du kinésithérapeute, un psychologue n'est pas un médecin spécialiste, mais il est également intéressant d'avoir cette information, car une consultation avec un psychologue pourrait être liée à des problèmes de santé mentale, qui sont également très fréquents chez les personnes en incapacité de travail.		
Variables liées aux hospitalisations					
Prestation C (hospitalisation)	Variable dérivée	Hospitalisation	Cette variable indique si la présentation est une hospitalisation.	Trimestriel 2006-2019	1 : la prestation est une hospitalisation 0 : autrement
Date d'admission de l'hospitalisation	ADMISSION – Date d'admission	Hospitalisation	Début du séjour hospitalier, 1 ^{er} jour d'hospitalisation. La date d'admission mentionnée dans la facturation d'une prestation hospitalière.	Mensuel 2006-2019	MM/AAAA
Admission de l'hospitalisation par une jour	FIRST_PREST	Hospitalisation	Pour les one day, le variable "date d'admission" est rarement remplie. Il faut alors employer cette variable.	Mensuel 2006-2019	MM/AAAA
Date de sortie de l'hôpital	DISCHARGE-Date de sortie de l'hôpital	Hospitalisation	La date de sortie de l'hôpital est celle reprise lors de la facturation d'une prestation hospitalière si elle est connue.	Mensuel 2006-2019	MM/AAAA
Nombre de jours d'hospitalisation	LOS – Length Of Stay (hospitalization)	Hospitalisation	Longueur du séjour (en jours) d'hospitalisation dans le trimestre de référence.	Trimestriel 2006-2019	Numéro

Données pseudonymisées issues du datawarehouse marché du travail et protection sociale de la BCSS

Nom de la variable	Code de la variable	Banque de données	Justification	Date et périodicité de les données	Valeurs possibles
Caractéristiques personnelles et socio-économiques					
Numéro d'identification de la sécurité sociale	INSZ	DWH_BCSS_RevenuDWH_BCSS_Revenu	Variable indispensable afin de regrouper les informations s'apparentant à un même individu.	Mensuel 2006-2019	Numéro codé
Date de naissance	D_geboor	DWH_RN_BCSS_BisTer	Variable à caractère personnel élémentaires comme variable de contrôle dans les estimations. Nous n'avons pas besoin de la date de naissance exacte de l'individu, mais nous avons besoin de l'âge au cours du trimestre. Il est important d'avoir l'âge comme variable continue car l'apparition de maladies pouvant conduire une personne à l'invalidité est étroitement et positivement corrélée avec l'âge. Il est également important de prendre en compte la distance qui sépare chaque individu de l'âge de pension	Trimestriel 2006-2019	MM/AAAA
Sexe	Geslacht	DWH_RN_BCSS_BisTer	Variable à caractère personnel élémentaires comme variable de contrôle dans les estimations. Le sexe est largement reconnu comme un facteur d'impact déterminant dans de nombreux phénomènes sociaux. Dans le cas de la santé et des incapacités au travail, plusieurs études démontrent ces différences.	Trimestriel 2006-2019	1 : homme 2 : femme
Catégorie d'études	Cat_études	<u>DWH_CREF_diplomes</u>	Variables permettant de différencier les effets en fonction du niveau et du type d'étude du demandeur d'emploi.	Trimestriel 2006-2019	codes
Première nationalité	Eerstre_natio	DWH_RN_Origine	Variable à caractère personnel élémentaires comme variable de contrôle dans les estimations. Le phénomène de l'immigration devient de plus en plus fréquent, notamment en Belgique. C'est pourquoi cette variable permet de savoir s'il existe des différences entre les immigrés et les nationaux, puisqu'ils sont également confrontés à des réalités socio-	Trimestriel 2006-2019	« Andere Europese landen » / « Autres Pays étrangers » / « Blgië » / « Congo, Burundi, Rwanda » / « Europe du nord » / « Europe du sud » / « Maghreb » / « Moyen Orient » / « Noord-Amerika » / « Oceanië » / « Onbekende vreemde

			économiques et d'intégration différentes qui peuvent avoir un impact sur la santé.		nationaliteit » / « Scandinavië » / « Zuid-en Midden-Amerika » / « overig Afrika » / « overig Azië » / missing
Nationalité	Natio	DWH_RN_B CSS_BisTer		Trimestriel 2006-2019	« Andere Europese landen » / « Autres Pays étrangers » / « Blgië » / « Congo, Burundi, Rwanda » / « Europe du nord » / « Europe du sud » / « Maghreb » / « Moyen Orient » / « Noord-Amerika » / Oceanië » / « Onbekende vreemde nationaliteit » / « Scandinavië » / « Zuid-en Midden-Amerika » / « overig Afrika » / « overig Azië » / missing
Domicile commune	Codnis	DWH_RN_B CSS_BisTer	Variable à caractère personnel élémentaires comme variable de contrôle dans les estimations. Il est important de prendre en compte les différences socio-économiques et géographiques dérivées de la commune de résidence. Le type de marché du travail (étendue et disponibilité), ainsi que l'offre de ressources en santé dans chaque lieu, peuvent entraîner des différences très importantes qui affectent notre domaine d'études. We need this variable at the arrondissement level.	Trimestriel 2006-2019	Codes (arrondissement)
Statut social	Sociale_stat	DWH_INAMI_CMI	Variable à caractère personnel élémentaires comme variable de contrôle dans les estimations. Variables nécessaires pour prendre en compte d'éventuels effets de revenus.	Trimestriel 2010-2019	2 : employée 3 : ouvrier 4 : indépendant 5 : conjoint aidant
Indicateur statistique	Aanwijz	DWH_INAMI_CMI	La même variable pour la période précédente.	Trimestriel 2006-2009	2 : employée 3 : ouvrier 4 : indépendant 5 : conjoint aidant
Régime	STELSEL	DWH_ONEm_Paiements	Variable à caractère personnel élémentaires comme variable de contrôle dans les estimations. Indique si la personne qui bénéficie de l'interruption de carrière est	Mensuel 2006-2019	Secteur privé / enseignement / secteur public

			employée dans le secteur public, le privé ou l'enseignement.		
Composition ménage	Samenst_g ezin	DWH_CIN	Variable de base nécessaire dans toute analyse sur l'incapacité de travail (afin de distinguer entre différents types de prestations selon le type d'individu).	Trimestriel 2006-2019	codes
Position lipro	LIPRO	DWH_RN_B CSS_BisTer	Variable à caractère personnel élémentaire comme variable de contrôle dans les estimations. C'est la position dans le ménage suivant la typologie LIPRO.	Trimestriel 2006-2019	Codes
Revenu annuel	inkomens	DWH_INAST I_RevenusIn dependants	Variable nécessaire pour prendre en compte d'éventuels effets de revenus.	Trimestriel 2006-2019	Par classes de 50€
décès	D_overl	DWH_RN_B CSS_BisTer_ Deces	Indique quand quelqu'un est mort, ce qui est utile pour voir la gravité des maladies.	Trimestriel 2006-2019	TT/AAAA
Données relatives à la position socio-économique					
Nomenclature de la position socio-économique	Nomenc	DWH_BCSS_ Nomenclatu reVarDer	Variables permettant d'identifier précisément le statut de chaque individu et d'observer les changements de statuts des individus entre différentes périodes.	Trimestriel 2006-2019	codes
En emploi	Variable dérivée		Variable indicatrice si la personne a été en pendant le trimestre. -- Indicateur qui dit si la personne est présente ou pas dans les fichiers DWH_ONSS_StatbaseDMFA / DWH_ONSSAPL_Statbase / DWH_ONSS_UniStatbaseDMFA. Cela détermine si quelqu'un est en emploi ou pas.	Trimestriel 2006-2019	1 : Oui 0 : No
Salaire	Sal100	DWH_ONSS_StatbaseDMFA	Le montant de la rémunération "ordinaire" du trimestre, c'est-à-dire toute rémunération brute (sauf les charges fiscales) qui ne constitue pas une indemnité de rupture, une prime ou un salaire d'attente (voir autres codes salariaux).	Trimestriel 2006-2016	Par classes de 50€
Salaire	Sal100	DWH_ONSS_UniStatbaseDMFA	La même variable pour la période précédente.	Trimestriel 2017-2019	Par classes de 50€

<u>Indépendant</u>	Variable dérivée de		Variable indicatrice si la personne est indépendante. -- Indicateur qui dit si la personne est présente ou pas dans le fichier DWH_INASTI_RGTI. Cela détermine si quelqu'un est indépendant ou pas.	Trimestriel 2006-2019	1 : Oui 0 : No
Revenu annuel	inkomens	DWH_INASTI_RevenusIndependants	Revenus reçus de l'INASTI. Cette variable est sur une base annuelle donc nous devons diviser par quatre.	Annuel 2006-2019	Par classes de 500€
Chômage		Variable dérivée	Variable indicatrice si la personne a connu une période de chômage pendant le trimestre. -- Si la personne perçoit des allocations de chômage de l'ONEm.	Trimestriel 2006-2019	1 : Oui 0 : No
Activation ONEm	Activation_ONEm	<u>DWH_BCSS_NomenclatureVarDer</u>	Position sur le marché du travail d'actif occupé en combinaison avec une activation de l'ONEm. On peut donc savoir si la personne perçoit des allocations de chômage de l'ONEm. Pour dérivée le variable "Chômage".	Trimestriel 2006-2019	1 : Oui 0 : No
Incapacité	Variable dérivée		Variable indicatrice si la personne a connu une période d'incapacité pendant le trimestre. Dérivée avec les variables suivantes.	Trimestriel 2006-2019	1 : Oui 0 : No
Donnés sur l'incapacité de travail					
Date de début invalidité	B_inval	DWH_INAMI_CMI	Cette variable indique la date de début en invalidité. Nous n'avons pas besoin de la date exacte, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.	Trimestriel 2006-2016	MM/AAAA
Date de début invalidité	b_inval	DWH_INAMI_CMI_RELAPSE	La même variable pour la période ultérieure	Trimestriel 2016-2019	MM/AAAA
Date de début période de paiement	Aanvangsdatum_beta	DWH_INAMI_Paiements	Cette variable indique la date de début du paiement, qu'elle puisse	Trimestriel 2006-2019	MM/AAAA

	lingsperiode		être différente du début en incapacité de travail, et il est important d'analyser cette différence, si elle existe. Nous n'avons pas besoin de la date exacte, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.		
Date de séance	D_zittin	DWH_INAMI_CMI	Cette variable indique la date de séance de l'invalidité. Nous n'avons pas besoin de la date exacte, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.	Trimestriel 2006-2019	MM/AAAA
Date début maladie (date début de l'incapacité)	Begindatum ziekte	DWH_INAMI_Paiements	Cette variable indique la date de début de la maladie, qu'elle puisse être différente du début en incapacité de travail, et il est important d'analyser cette différence, si elle existe. Nous n'avons pas besoin de la date exacte, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.	Trimestriel 2006-2019	MM/AAAA
Début maladie (date début incapacité de travail primaire)	B_ziekte	DWH_INAMI_CMI		Trimestriel 2006-2019	MM/AAAA
Décision conseil médical de l'invalidité	Gri_cod	DWH_INAMI_CMI	L'invalidité est reconnue par le Conseil médical de l'invalidité sur la base d'un rapport détaillé établi par le médecin-conseil de l'organisme assureur. Le Conseil doit prendre une décision dans les quatre dernières semaines de la période d'incapacité de travail primaire. Cette information est pertinente pour la présente étude	Trimestriel 2006-2019	codes
Jours indemnisés	Vergoede_dagen	DWH_INAMI_Paiements	Variable nécessaire pour connaître l'importance/ la durée de l'invalidité et ainsi, distinguer entre différents types d'invalidité (courte	Trimestriel 2006-2019	métrique

			durées >< longue durée, occurrences multiples, etc.).		
Montant indemnité (Invalidité)	bedrag	DWH_INAMI_Paiements	Variable nécessaire pour prendre en compte d'éventuels effets de revenus (invalidité).	Trimestriel 2006-2019	métrique
R_exclus			Variable permettant d'identifier le statut exact de l'individu à la fin du trimestre.	Trimestriel 2006-2019	
Date de début incapacité de travail	Begindatum	DWH_CIN	Cette variable indique la date de début en Incapacité primaire. Nous n'avons pas besoin de la date exacte, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.	Mensuel 2006-2019	MM/AAAA
Date de fin incapacité de travail	Einddatum	DWH_CIN	Cette variable indique la date de séance en Incapacité primaire. Nous n'avons pas besoin de la date exacte, seulement au trimestre de l'année, puisque nous travaillons avec des données trimestrielles. De cette façon, il est impossible d'identifier le moment exact où la personne reçoit la prestation et il suffit de mener notre étude.	Mensuel 2006-2019	MM/AAAA
Montant (Incapacité de travail primaire)	Bedrag	DWH_CIN	Variable nécessaire pour prendre en compte d'éventuels effets de revenus (incapacité de travail primaire).	Mensuel 2006-2019	Métrique (Montant de l'indemnité en classes de 50€)
Nature de l'indemnité	Aard_uitkering	DWH_CIN	Variable permettant de distinguer les effets en fonction de différents types d'indemnités.	Mensuel 2006-2019	codes
Nombre de jours d'incapacité de travail	Aantal_dagen_ao	DWH_CIN	Variable nécessaire pour connaître l'importance/ la durée de l'incapacité et ainsi, distinguer entre différents types d'incapacités (courte durées >< longue durée, occurrences multiples, etc.).	Mensuel 2006-2019	métrique
R_Exclus	R_Exclus	DWH_CIN	Variable permettant d'identifier le statut exact de l'individu à la fin du trimestre.	Mensuel 2006-2019	
Type d'indemnité	Srt_uitkering	DWH_CIN	Variables permettant de distinguer les effets entre différents types de bénéficiaires de l'allocation d'incapacité.	Mensuel 2006-2019	Pas d'indemnité / indemnité / indemnité pour travail adapté
Type de jours	Srt_dagen	DWH_CIN		Mensuel 2006-2015	Incapacité de travail primaire / Ecartement du travail pour cause de grossesse / congé de maternité / Ecartement

					du travail pour cause d'allaitement (congé d'allaitement prophylactique) / ...
Code indemnité	Betalingscode	DWH_INAMI_Paiements	Variables permettant de distinguer les effets entre différents types de bénéficiaires de l'allocation d'invalidité, ainsi qu'entre différents maladies et motifs de sortie du statut.	Trimestriel 2006-2017	codes
Code médical	Gr_ziekt	DWH_INAMI_CMI		Trimestriel 2006-2017	Codes (maladies)
Code sortie	Uittredingscodes	DWH_INAMI_Paiements		Trimestriel 2006-2017	Reprise / décès/ mise a la retraite / mutation /suspension du paiement... / renonciation aux indemnités / toutes autres causes / pas d'application
Régime	Stelsel	DWH_INAMI_Paiements		Trimestriel 2006-2019	Salarie/ Independent
Données sur l'emploi					
Date de début	BEGINACT	DWH_ONE m_Paiements	Date de début de l'emploi dans le cadre d'une mesure d'activation. Variable nécessaire pour identifier le régime auquel chaque individu est soumis en fonction des évolutions législatives.	Mensuel 2006-2019	MM/AAAA
Contrat de travail	CONTRAT	DWH_ONE m_Paiements	Permet d'identifier le statut des individus ayant souscrit à un contrat de d'interruption de carrière ou de crédit-temps.	Mensuel 2006-2019	
Classe de travailleur	clatra	DWH_ONSS_StatbaseD MFA	Variables permettant d'obtenir des résultats se basant sur le parcours professionnel détaillé de l'individu au cours d'une période donnée. Ces variables permettent également de distinguer les effets entre différentes situations professionnelles actuelles ou passées de l'individu. Certaines de variables (notamment la commission paritaire applicable) permet d'identifier le respect des conditions pour bénéficier de certaines allocations (notamment en ce qui concerne les prépensions). D'autres variables (telles que la classe salariale ou le salaire journalier) permettent de prendre	Trimestriel 2006-2016	1 : ouvriers 2 : employés 3 : fonctionnaires 4 : autres

			en compte d'éventuels effets de revenus.		
Classe de travailleur	clatra	DWH_ONSS _UniStatbas eDMFA	La même variable pour la période ultérieure.	Trimestriel 2017- 2019	1 : ouvriers 2 : employés 3 : fonctionnaires 4 : autres
Classe de travailleur spéciale	Clatr2	DWH_ONSS _StatbaseD MFA	Variables permettant d'obtenir des résultats se basant sur le parcours professionnel détaillé de l'individu au cours d'une période donnée. Ces variables permettent également de distinguer les effets entre différentes situations professionnelles actuelles ou passées de l'individu. Certaines de variables (notamment la commission paritaire applicable) permet d'identifier le respect des conditions pour bénéficier de certaines allocations (notamment en ce qui concerne les prépensions). D'autres variables (telles que la classe salariale ou le salaire journalier) permettent de prendre en compte d'éventuels effets de revenus.	Trimestriel 2006-2016	codes
Classe de travailleur spéciale	Clatr2	DWH_ONSS _UniStatbas eDMFA	La même variable pour la période ultérieure.	Trimestriel 2017-2019	codes
Classe salariale	Salcla	DWH_ONSS _StatbaseD MFA	Variables permettant d'obtenir des résultats se basant sur le parcours professionnel détaillé de l'individu au cours d'une période donnée. Ces variables permettent également de distinguer les effets entre différentes situations professionnelles actuelles ou passées de l'individu. Certaines de variables (notamment la commission paritaire applicable) permet d'identifier le respect des conditions pour bénéficier de certaines allocations (notamment en ce qui concerne les prépensions). D'autres variables (telles que la classe salariale ou le salaire journalier) permettent de prendre en compte d'éventuels effets de revenus.	Trimestriel 2006-2016	Classes

Classe salariale	salcla	DWH_ONSS_ UniStatbas_eDMFA	la même variable pour la période ultérieure	Trimestriel 2017 – 2019	Classes
Classe temps partiel	Catept	DWH_ONSS_StatbaseDMFA	Variables permettant d'obtenir des résultats se basant sur le parcours professionnel détaillé de l'individu au cours d'une période donnée. Ces variables permettent également de distinguer les effets entre différentes situations professionnelles actuelles ou passées de l'individu. Certaines de variables (notamment la commission paritaire applicable) permet d'identifier le respect des conditions pour bénéficier de certaines allocations (notamment en ce qui concerne les prépensions). D'autres variables (telles que la classe salariale ou le salaire journalier) permettent de prendre en compte d'éventuels effets de revenus.	Trimestriel 2006-2016	0% 0-15% 16-35% 36-65% 66-95% 96-99% >99%
Classe temps partiel	catept	DWH_ONSS_ UniStatbas_eDMFA	la même variable pour la période ultérieure	Trimestriel 2017 - 2019	0% 0-15% 16-35% 36-65% 66-95% 96-99% >99%
Commission paritaire	Nr_cp	DWH_ONSS_StatbaseDMFA	Indique le numéro de la commission paritaire, qui est déterminé en fonction de l'activité économique de l'entreprise. Ce qui pourrait sans doute être très important dans la conduite d'une incapacité de travail.	Trimestriel 2006-2016	codes
Commission paritaire	Nr_cp	DWH_ONSS_ UniStatbas_eDMFA	La même variable pour la période ultérieure.	Trimestriel 2017-2019	codes
Jours assimilés	Jrsasm	DWH_ONSS_StatbaseDMFA	Cette variable fait référence au cheminement de travail de chaque individu, ce qui pourrait également avoir un effet à la fois sur la détérioration de la santé de la personne et sur la décision d'amorcer une démarche pour demander une invalidité.	Trimestriel 2006-2016	Métrique (nombre total des jours assimilés du trimestre)

Jours assimilés	Jrsasm	DWH_ONSS_ UniStatbas eDMFA	la même variable pour la période ultérieure	Trimestriel 2017-2019	Métrique (nombre total des jours assimilés du trimestre)
Pourcentage temps partiel	Tauxpt	DWH_ONSS_ StatbaseD MFA	Le type de journée de travail et le nombre d'heures travaillées pourrait avoir un effet à la fois sur la détérioration de la santé de la personne (burnout, problèmes de santé mentale et musculo-squelettiques) et sur la décision d'amorcer une démarche pour demander une invalidité.	Trimestriel 2006-2016	Métrique (le pourcentage d'un travail à temps plein que le travailleur à temps partiel effectue)
Pourcentage temps partiel	Tauxpt	DWH_ONSS_ UniStatbas eDMFA	la même variable pour la période ultérieure	Trimestriel 2017-2019	Métrique (le pourcentage d'un travail à temps plein que le travailleur à temps partiel effectue)
Salaire journalier	Saljrs	DWH_ONSS_ StatbaseD MFA	Le salaire reflète le pouvoir d'achat de l'individu, information nécessaire à contrôler dans les études économétriques à caractère socio-économique, puisqu'il peut générer des disparités importantes.	Trimestriel 2006-2016	Métrique (le salaire journalier moyenne)
Salaire journalier	Saljrs	DWH_ONSS_ UniStatbas eDMFA	La même variable pour la période ultérieure.	Trimestriel 20017-2019	Métrique (le salaire journalier moyenne par classes de 1€)
Code qui indique si le travailleur a été occupé via le régime des titres-services.	Proximity obs	<u>DWH_ONSS_ StatbasePI us</u>	Il a été démontré que les travailleurs du SVS sont particulièrement susceptibles d'avoir une invalidité et qu'ils constituent donc un sous-groupe important à étudier lorsqu'on essaie de prédire les modèles d'entrées en incapacité de travail.		
Code nace	codnac	DWH_ONSS_ APL_ Statplu s_PPO	Indique le secteur dans lequel une personne était employée, ce qui pourrait sans doute être très important dans la conduite d'une incapacité de travail.	Trimestriel 2006-2007	Codes (premiers 3 chiffres)
Code nace	codnac	DWHONSS_ StatbaseDM FA	La même variable pour la période ultérieure.	Trimestriel 2008-2016	Codes (premiers 3 chiffres)
Code nace	codnac	DIWONSS_ U niStatbaseD MFA	La même variable pour la période ultérieure.	Trimestriel 2017-2019	Codes (premiers 3 chiffres)

Équivalent temps plein, journées assimilées incluses	tauxaa	DWH_ONSS APL_Statbas eDMFA	L'équivalent temps plein donne une indication du degré d'occupation d'un emploi au cours du trimestre. Ce qui pourrait sans doute être très important dans la conduite d'une incapacité de travail.	Trimestriel 2006-2016	Métrique
Équivalent temps plein, journées assimilées incluses	tauxaa	DWH_ONSS APL_UniStat baseDMFA	La même variable pour la période ultérieure.	Trimestriel 2017-2019	Métrique