

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

CSI/CSSS/24/060

**DÉLIBÉRATION N° 24/036 DU 5 MARS 2024 PORTANT SUR LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES RELATIVES À LA SANTÉ PAR L'AGENCE INTERMUTUALISTE, LE SERVICE EXTERNE POUR LA PRÉVENTION ET LA PROTECTION AU TRAVAIL IDEWE, LE SERVICE PUBLIC FÉDÉRAL SANTÉ PUBLIQUE ET LA BANQUE CARREFOUR DE LA SÉCURITÉ SOCIALE À LA VRIJE UNIVERSITEIT BRUSSEL DANS LE CADRE DE SON PROJET DE RECHERCHE « HEARTWISE : THE HEALTH ECONOMIC ANALYSIS OF RETURN TO WORK AND THE SOCIETAL COST OF CARDIOVASCULAR DISEASES »**

Le Comité de sécurité de l'information, chambre sécurité sociale et santé (dénommé ci-après « le Comité ») ;

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

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

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 du groupe de recherche « Interuniversity Centre for Health Economics Research » (I-CHER) de la Vrije Universiteit Brussel ;

Vu le rapport d'auditorat de la Plate-forme eHealth du 15 février 2024 ;

Vu le rapport de monsieur Michel Deneyer ;

Émet, après délibération, la décision suivante, le 5 mars 2024 :

**A. OBJET DE LA DEMANDE**

1. Le groupe de recherche I-CHER a introduit une demande auprès du Comité de sécurité de l'information afin de pouvoir obtenir des données à caractère personnel pseudonymisées relatives à la santé de la part de l'Agence intermutualiste, de la Banque Carrefour de la sécurité sociale, du service externe pour la prévention et la protection au travail (IDEWE) et du SPF Santé publique dans le cadre de son projet de recherche « HeaRTWise : The health economic analysis of return to work and the societal cost of cardiovascular diseases ».
2. Ce projet a pour objectif principal de déterminer le coût sociétal<sup>1</sup> des maladies cardiovasculaires (MCV). Il vise à améliorer les pronostics du patient dans le domaine professionnel suite à une maladie cardiovasculaire. Par ailleurs, les chercheurs tentent aussi de prédire l'emploi après une MCV au niveau de la population.

### **Procédure de sélection**

3. Le groupe-cible principal de cette étude sont les personnes ayant été hospitalisées en raison d'une maladie cardiovasculaire (MCV) comme diagnostic principal entre le 1<sup>er</sup> janvier 2017 et le 31 décembre 2018. L'étude tâche de retracer le trajet de ces personnes sur une période d'un an avant la date indice jusque deux ans après la date indice<sup>2</sup>. La date indice est déterminée sur la base de la première hospitalisation<sup>3</sup>.
4. Les diagnostics MCV sont obtenus auprès du SPF Santé publique<sup>4</sup>. Dans le cadre de cette étude, une sélection des maladies graves les plus fréquentes du système cardiovasculaire est effectuée et le but étant de dresser la carte de l'emploi et du coût sociétal. De manière spécifique, l'étude se focalise sur les personnes ayant reçu un des diagnostics suivants comme diagnostic principal vérifié lors de l'hospitalisation, sur la base de la classification ICD-10-BE :
  - Maladies cardiaques ischémiques I20-I25
  - Autres maladies cardiaques
    - I30: Péricardite aiguë
    - I33-36: Endocardite aiguë et subaiguë, affections non rhumatismales de la valve mitrale, aortique et tricuspide.
    - I38: Endocardite
    - I40: Myocardite aiguë
    - I46.0: Arrêt cardiaque réanimé avec succès
    - I48: Fibrillation et flutter auriculaires
    - I49.0: Fibrillation et flutter ventriculaires
    - I50: Insuffisance cardiaque
  - Maladies cérébrovasculaires :
    - I60-64: Hémorragie sous-arachnoïdienne, hémorragies intracérébrales, autres hémorragies intracrâniennes, infarctus cérébral, AVC.
  - Maladie des artères :

---

<sup>1</sup> Les frais médicaux directs et les frais indirects en raison de la perte de productivité par rapport à la consommation de soins de base 1 an avant la maladie.

<sup>2</sup> Date indice – 365 jours jusque date indice + 730 jours (366 jours si année bissextile).

<sup>3</sup> En raison d'une MCV comme diagnostic principal entre le 01/01/2017 et le 31/12/2018.

<sup>4</sup> Variable: "A2\_CODE\_DIAG\_VERIF\_ADM".

- I70: Anévrisme et dissection aortique.

La date de la première hospitalisation en raison d'une MCV comme diagnostic principal (variable : « date d'incidence ») est considérée comme la « date indice ». Si un patient a été hospitalisé à la fois en 2017 et en 2018 en raison d'une MCV, seule la date de la première hospitalisation est considérée comme la date indice. Les autres variables basées sur une date ou une durée sont toujours exprimées en « nombre de jours par rapport à la date indice ».

5. Pour cette étude, l'ensemble de la population des personnes ayant été hospitalisées en raison d'une MCV en 2017 (n= 283.865) et en 2018 (n= 285.224) a été incluse<sup>5</sup>. Etant donné que l'étude vise également à prédire l'emploi après une MCV au niveau de la population, il est nécessaire que les chercheurs effectuent une sélection dans leurs données de tous les patients belges en âge actif ayant été hospitalisés en raison d'un diagnostic MCV en 2017 et 2018. Au sein de la population active (18-65 ans), environ 130.000 hospitalisations pour MCV au sens large ont été enregistrées en 2017.
6. Le flux de données est repris ci-dessous.

## **II. COMPÉTENCE**

7. Conformément à l'article 42, § 2, 3°, de la loi du 13 décembre 2006 *portant dispositions diverses en matière de santé*, toute communication de données à caractère personnel relatives à la santé requiert, sauf les exceptions prévues, une délibération de principe de la chambre sécurité sociale et santé du Comité de sécurité de l'information.
8. Le Comité estime dès lors qu'il est compétent pour se prononcer sur cette communication de données à caractère personnel relatives à la santé.

## **III. EXAMEN DE LA DEMANDE**

### **A. ADMISSIBILITÉ**

9. Le traitement de données à caractère personnel relatives à la santé est en principe interdit, et ce conformément au prescrit de l'article 9, §1<sup>er</sup> 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 (RGPD).
10. L'interdiction du traitement de données à caractère personnel relatives à la santé ne s'applique pas lorsque le traitement est nécessaire à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques, conformément à l'article 89, paragraphe 1, sur la base du droit de l'Union ou du droit d'un État membre, qui doit être proportionné à l'objectif poursuivi, respecter l'essence du droit à

---

<sup>5</sup> SPF Santé publique, communication personnelle du 6 février 2023.

la protection des données et prévoir des mesures appropriées et spécifiques pour la sauvegarde des droits fondamentaux et des intérêts de la personne concernée<sup>6</sup>.

11. Le Comité prend acte du fait que le Comité d'éthique de l'UZ Brussel a formulé un avis positif pour cette étude.
12. L'article 279 de la loi-programme (I) du 24 décembre 2002 stipule que toute transmission de données à caractère personnel à l'AIM requiert une autorisation du Comité sectoriel de la sécurité sociale, visé à l'article 37 de la loi du 15 janvier 1990 relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale.
13. 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 précise que la Banque Carrefour de la sécurité sociale recueille des données sociales, 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.
14. L'article 10 de l'arrêté royal du 27 avril 2007<sup>7</sup> prévoit que les données qui sont reprises dans la base de données hospitalières puissent être mises à la disposition de tiers dans le cadre d'une étude unique et temporaire. Ces études doivent s'inscrire dans le cadre des objectifs visés à l'article 3 et 19 de cet arrêté. En outre, l'étude doit toujours être de nature purement scientifique et donc ne poursuivre aucun but commercial.
15. Le service externe pour la prévention et la protection au travail IDEWE exécute ses tâches conformément à la loi sur le bien-être du 4 août 1996, au Code du bien-être au travail (plus précisément le Livre I, titre 4 Mesures relatives à la surveillance de la santé des travailleurs, I.4-2d en I.4-83), à la loi du 22 août 2002 relative aux droits du patient, au RGPD, à la loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel, au contrat de services avec le client et aux règles disciplinaires de l'Ordre des médecins, lorsqu'il s'agit d'informations médicales.
16. A la lumière de ce qui précède, le Comité est d'avis qu'il existe un fondement admissible pour le traitement des données à caractère personnel relatives à la santé envisagé.

## **B. PRINCIPES RELATIFS AU TRAITEMENT DE DONNÉES À CARACTÈRE PERSONNEL**

### **1. FINALITÉS**

17. En vertu de 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.

---

<sup>6</sup> Art. 9, alinéa 2, j) du RGPD.

<sup>7</sup> Arrêté royal du 27 avril 2007 déterminant les règles suivant lesquelles certaines données hospitalières doivent être communiquées au Ministre qui a la Santé publique dans ses attributions.

18. Les maladies cardiovasculaires (MCV) sont mondialement la principale cause de décès et la « charge » des MCV se situe à trois niveaux avec un taux de mortalité élevé, une morbidité considérable et un coût financier élevé. Les statistiques européennes en matière de MCV estiment le coût total des maladies cardiaques ischémiques, des maladies cardiovasculaires et cérébrovasculaires en 2015 à 6,845 milliards d'euros.
19. Diverses études épidémiologiques ont démontré ces dernières années que l'incidence et la mortalité des MCV a diminué dans les pays développés, ce qui est mis sur le compte d'une meilleure maîtrise des facteurs de risque, de la promotion de meilleurs soins, de traitements plus effectifs et de l'augmentation des budgets pour les services de santé. Grâce aux progrès réalisés au niveau du traitement médical, le nombre de patients rétablis (en partie) a augmenté pour ces maladies. Ces développements dans les soins de santé vont de pair avec d'autres évolutions sociales (augmentation de l'âge de la retraite) qui font que le nombre de personnes atteintes d'une affection a augmenté au sein de la population active. Au sein de la population active (18-65 ans) en Belgique, environ 130.000 hospitalisations en raison de MCV (codes ICD-10 I00-I99) ont été enregistrées en Belgique en 2017. Par ailleurs, l'incidence de MCV dans le groupe des personnes plus âgées a certes diminué, mais une diminution similaire n'apparaît pas dans le groupe des personnes plus jeunes.
20. Ces évolutions tendent à indiquer que la charge économique et la charge de santé des MCV augmenteront de manière dramatique dans les décennies à venir. D'un point de vue économique et du point de vue du patient, la réintégration sur le marché du travail revêt donc une importance cruciale. L'invalidité suite à un AVC (par exemple) et d'autres MCV a des conséquences négatives considérables pour la qualité de vie individuelle, la productivité au travail et la charge économique des MCV. Pour réduire la charge personnelle, sociétale et économique suite à l'incapacité de travail, il est essentiel de faciliter la reprise du travail. Il existe un rapport positif entre la reprise du travail après une MCV et le pronostic clinique du patient et sa qualité de vie. Pour faciliter la reprise du travail, il est évidemment important de connaître les facteurs qui ont une influence sur le retour au travail.

Même si des études ont déjà démontré que la réadaptation fonctionnelle, la réintégration sur le marché du travail et la prévention secondaire sont des composants essentiels d'une gestion réussie de la maladie, il existe encore beaucoup d'incertitude quant à l'ensemble des facteurs permettant de prédire la reprise du travail après une MCV. En identifiant les prédicteurs d'un retour au travail et les patients avec une grande probabilité de ne pas reprendre le travail, il est possible d'attribuer de manière plus efficace l'aide financière et les ressources destinées aux programmes de réadaptation et de réintégration du marché du travail au profit des patients avec un profil à risque élevé. Des prédictions précises quant aux résultats au niveau professionnel après une MCV peuvent par ailleurs offrir un pronostic plus réaliste au patient, à ses proches et aux professionnels de la santé et contribuer ainsi à fixer des objectifs appropriés dans le cadre de la réadaptation.

21. Des études antérieures ont permis d'identifier différents groupes complexes de variables comprenant les prédicteurs les plus pertinents en matière de reprise du travail après une MCV : des variables liées à la santé, des variables socio-démographiques et des variables

psychosociales liées au travail. Pour développer un modèle de prédiction en matière de reprise du travail sur la base de cet ensemble complexe de facteurs, il est nécessaire de disposer d'une vaste série de données cliniques détaillées, or le développement d'une telle banque de données demande beaucoup de temps et de ressources. L'utilisation de données administratives, dérivées de sources administratives alimentées de manière routinière, offre une alternative qui permet de gagner du temps et de l'argent. Pour développer de manière moins onéreuse des modèles de prédiction précis pour la population belge de patients atteints de MCV, les chercheurs pourraient avoir recours aux sources de données administratives belges existantes.

Des données de santé et des données socio-démographiques sont présentes dans les banques de données de population administratives du SPF Santé publique, de l'Agence intermutualiste et de la Banque Carrefour de la sécurité sociale. Pour une partie de la population enregistrée dans les sources de données administratives, les chercheurs peuvent également disposer de données liées au travail des services de santé et de sécurité au travail tels qu'IDEWE. Une prédiction précise de l'emploi après une MCV requiert un couplage de toutes ces sources de données, ce qui aboutit à un vaste ensemble combiné de données au niveau de la population. Le couplage de ces données offre une valeur ajoutée d'un point de vue scientifique puisque ceci permet de réaliser des études au niveau de la population (étude de population). Par ailleurs, cette étude pourra contribuer à de futures études du rapport coût/efficacité en ce qui concerne le développement de nouvelles stratégies préventives visant à diminuer la charge de morbidité pour cette population.

- 22.** Ce projet se focalise sur (les prédictions de) l'emploi de la population belge de patients MCV et du coût sociétal y afférent. Ci-après figurent quelques sous-objectifs :
- 1° Emploi (reprise du travail) après MCV
    - a. Analyse de l'emploi après l'hospitalisation indice ; hospitalisation en raison de MCV entre le 01/01/2017 et le 31/12/2018 après 1 an et après 2 ans.
    - b. Couplage de données (AIM-SPF-BCSS-IDEWE)
    - c. Développement d'un modèle de prédiction précis en matière de reprise du travail
    - d. Analyse de variables prédictives en matière d'emploi après MCV.
  - 2° Coût sociétal des MCV.
    - a. Analyse du coût sociétal (direct + indirect) de MCV dans la période 2017-2020 et la consommation de soins de base 1 an avant l'hospitalisation indice.
    - b. Etude des variables prédictives pour le coût sociétal de MCV.
- 23.** Compte tenu des finalités du traitement décrites ci-avant, 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. MINIMISATION DES DONNÉES**

- 24.** Selon l'article 5 du RGPD, 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.

25. Il s'agit d'un traitement ultérieur. Le but initial pour l'Agence intermutualiste est d'une part la collecte des données en matière de consommation de soins et de coûts via les différentes mutualités en fonction du remboursement par l'assurance obligatoire soins de santé et d'autre part la recherche scientifique (d'appui à la politique). Le SPF Santé publique recueille des résumés hospitaliers minimums (RHM), notamment en ce qui concerne la consommation de soins et les coûts dans les hôpitaux dans le cadre du financement des hôpitaux et de la recherche scientifique (d'appui à la politique). La Banque Carrefour de la sécurité sociale enregistre des données relatives à la sécurité sociale notamment en vue d'inventorier les cotisations sociales et les allocations. IDEWE enregistre des données (psychosociales) en relation avec le travail et des données relatives à l'intégration sur le marché du travail dans le cadre de l'emploi après une maladie cardiovasculaire.
26. Une liste des données demandées ainsi qu'une motivation circonstanciée par variable sont jointes ci-après.

### **3. LIMITATION DE LA CONSERVATION**

27. Selon l'article 5, §1<sup>er</sup>, e) du RGPD, 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; les données à caractère personnel peuvent être conservées pour des durées plus longues dans la mesure où elles seront traitées exclusivement à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques conformément à l'article 89, §1<sup>er</sup>, pour autant que soient mises en œuvre les mesures techniques et organisationnelles appropriées requises par le règlement afin de garantir les droits et libertés de la personne concernée (limitation de la conservation).
28. La période étudiée débute deux années avant la date d'incidence<sup>8</sup> et se termine une année après la date d'incidence. Dans le cadre du suivi scientifique du projet (par exemple pour le contrôle des données statistiques suite à des publications scientifiques à comité de lecture), l'ensemble des données sera conservé jusque quatre ans après la transmission des données. Les clés de conversion pour la pseudonymisation des numéros de registre national sont conservées pendant dix ans par le fournisseur de données dans le cadre d'éventuelles demandes de données futures pour des études de suivi.
29. Il est nécessaire de conserver le lien entre la série de données pseudonymisées et l'identité des personnes. Le TTP eHealth conserve ce lien. Pour permettre le couplage des données entre les divers fournisseurs<sup>9</sup>, il est important que le lien entre la série de données pseudonymisées et l'identité des personnes ne se perde pas.
30. Le Comité estime que le délai de conservation est raisonnable.

### **4. TRANSPARENCE**

---

<sup>8</sup> Diagnostic principal de MCV entre le 01/01/2017 et le 31/12/2018.

<sup>9</sup> SPF Santé publique, BCSS, IDEWE et AIM.

31. Lorsque les données à caractère personnel n'ont pas été collectées auprès de la personne concernée, le responsable du traitement est tenu de fournir les informations visées à l'article 14, §§1 et 2 du RGPD à la personne concernée.
32. Cette obligation ne s'applique cependant pas lorsque la fourniture 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, alinéa 1<sup>er</sup>, ou dans la mesure où l'obligation visée à l'alinéa 1<sup>er</sup> du présent article est susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs dudit traitement. En pareils cas, le responsable du traitement prend des mesures appropriées pour protéger les droits et libertés ainsi que les intérêts légitimes de la personne concernée, y compris en rendant les informations publiquement disponibles.
33. Les données constituent des données pseudonymisées qui, par définition, en l'absence de données relatives à l'identité, ne permettent pas d'identifier directement les personnes concernées. Par ailleurs, les chercheurs n'ont pas accès aux données à caractère personnel des intéressés et la série de données est tellement vaste (environ 550.000 patients) que la communication d'informations à tous les intéressés n'est pas faisable, que ce soit d'un point de vue pratique que d'un point de vue financier.
34. Le Comité est d'avis qu'il existe suffisamment de transparence quant au traitement envisagé.

## **5. SÉCURITÉ DE L'INFORMATION**

35. Selon l'article 5, §1<sup>er</sup>, 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).
36. Le Comité constate qu'une analyse d'impact relative à la protection des données a eu lieu, conformément à l'article 35 du RGPD, et qu'elle lui a été transmise.
37. Le Comité constate que le Centre fédéral d'expertise des soins de santé (KCE) réalisera une analyse de risque « small cell » avant la transmission des données aux chercheurs.
38. Le Comité constate que la Vrije Universiteit Brussel a désigné un médecin responsable du traitement des données à caractère personnel relatives à la santé, ainsi qu'un délégué à la protection des données.
39. Le Comité constate que les collaborateurs de la Vrije Universiteit Brussel ont un devoir de confidentialité contractuel par rapport aux données qu'ils traitent dans le cadre de leur fonction.



40. Le Comité constate que la Plate-forme eHealth intervient comme tiers de confiance (TTP) pour le couplage et la pseudonymisation des données.
41. Le Comité rappelle que ni la Vrije Universiteit Brussel ni ses collaborateurs ne peuvent entreprendre des démarches pour réidentifier les personnes concernées. Les résultats de l'étude doivent être publiés sous forme anonyme.
42. Le Comité 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.
43. Le Comité estime nécessaire de rappeler que depuis le 25 mai 2018, la Vrije Universiteit Brussel est tenue de respecter les dispositions et les principes 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 (règlement général sur la protection des données). Ces instances sont également tenues de respecter les dispositions de la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel*.

Par ces motifs,

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

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 des données 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.

La présente délibération entre en vigueur le 20 mars 2024.

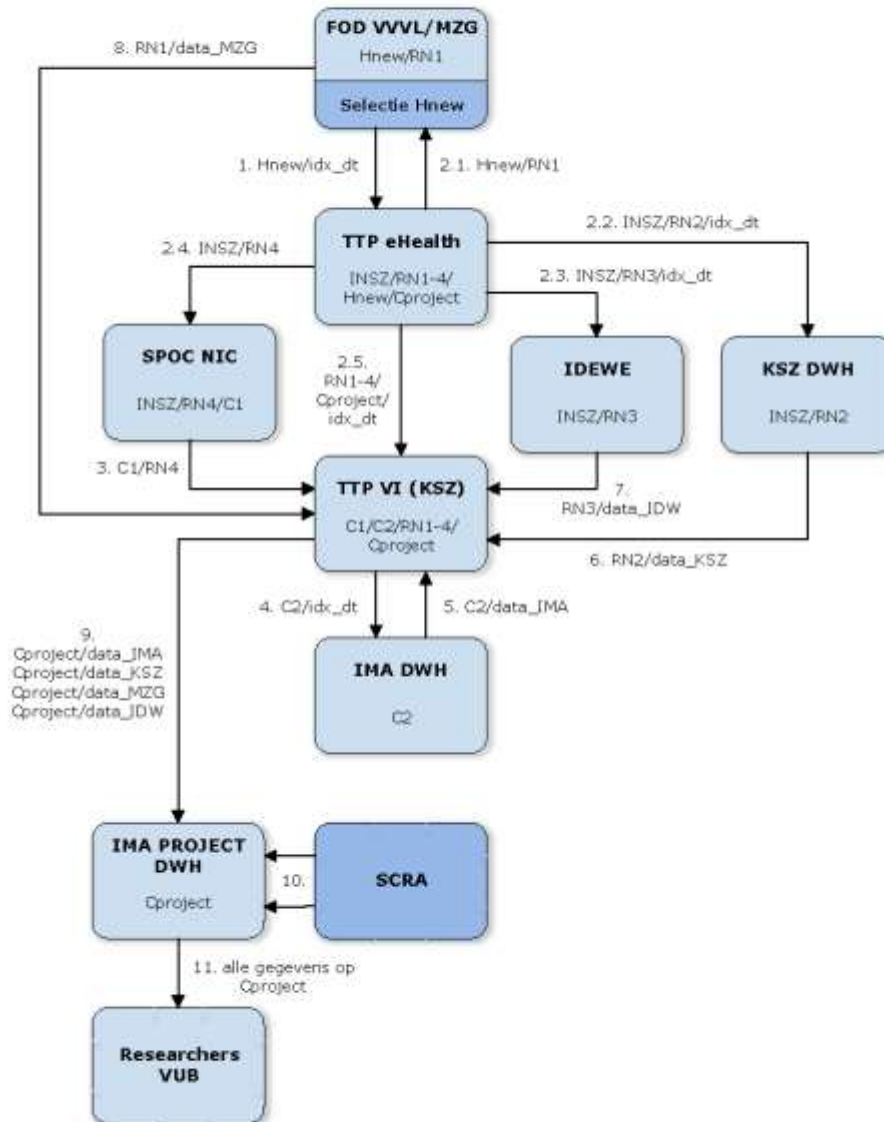
Michel DENEYER  
Président

Le siège de la chambre sécurité sociale et de la 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).

**Bijlage 1: overzicht gegevensstroom**  
**1. Schematisch overzicht gegevensstromen**



Gegevensstroom koppeling IDEWE, IMA, KSZ & FOD VVVL data:  
 HeaRTWise studie (IMA referentie: P\_HRTW01)



*Figuur 1: Schematische voorstelling van data-extracties*

## 2. Beschrijving van de gegevensstromen en de codering

(IMA referentie: P\_HRTW01)

In de context van deze studie selecteert de verantwoordelijke van de minimale ziekenhuisgegevens bij de Federale Overheidsdienst Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu (FOD VVVL/MZG) de Hnew (de identifier die eHealth standaard gebruikt in de MZG flow sinds 2012) van alle personen met een ziekenhuisopname ten gevolge van een hart- en vaatziekte tussen 01/01/2017 en 31/12/2018. Voor elke geselecteerde persoon wordt tevens de (eerste) incidentiedatum toegevoegd (indexdatum oftewel idx\_dt). Hierna wordt de persoonslijst met datums (Hnew/idx\_dt) overgebracht naar de Trusted Third Party (TTP) eHealth (**stap 1**).

De TTP eHealth converteert elke Hnew naar het identificatienummer van de sociale zekerheid (INSZ) en maakt hierna voor elke INSZ in de lijst een unieke Cproject aan, samen met 4 random numbers: RN1, RN2, RN3 en RN4. Hierna zal de TTP eHealth de volgende stappen ondernemen:

- De conversielijst Hnew/RN1 wordt overgebracht naar de FOD VVVL/MZG (**stap 2.1**).
- De conversielijst INSZ/RN2 inclusief geassocieerde indexdatums idx\_dt wordt overgebracht naar de Kruispuntbank van de Sociale Zekerheid (KSZ) (**stap 2.2**).
- De conversielijst INSZ/RN3 inclusief geassocieerde indexdatums idx\_dt wordt overgebracht naar de externe dienst voor preventie en bescherming op het werk IDEWE (**stap 2.3**).
- De conversielijst INSZ/RN4 wordt overgebracht naar het uniek aanspreekpunt van het Nationaal Intermutualistisch College (SPOC NIC) (**stap 2.4**).
- De conversielijst Cproject/RN1/RN2/RN3/RN4 inclusief indexdatums wordt overgebracht naar de Trusted Third Party van de verzekeringsinstellingen – Kruispuntbank Sociale Zekerheid (TTP VI (KSZ)) (**stap 2.5**).

SPOC NIC converteert in de INSZ/RN4 lijst elke INSZ naar de code van de verzekeringsinstellingen (C1) en brengt de nieuwe lijst C1/RN4 over naar de TTP VI (KSZ) (**stap 3**).

De TTP VI (KSZ) converteert elke C1 uit de C1/RN4 lijst naar de code C2 van het InterMutualistisch Agentschap (IMA) en plaatst de C2 codes inclusief geassocieerde indexdatums (C2/idx\_dt) in het IMA Data Warehouse (DWH) (**stap 4**).

Het IMA selecteert alle nodige gegevens op C2 (data\_IMA) en transformeert alle absolute datums naar aantal kalenderdagen ten opzichte van de indexdatum van elke patiënt. Hierna worden de gegevens overgebracht naar de TTP VI (KSZ) (**stap 5**).

De KSZ selecteert uit het KSZ DWH alle nodige gegevens op INSZ (data\_KSZ) en transformeert alle absolute datums naar aantal kalenderdagen ten opzichte van de indexdatum van elke patiënt. Hierna wordt elke INSZ vervangen door de corresponderende RN2 en worden de gegevens overgebracht naar de TTP VI (KSZ) (**stap 6**).

IDEWE selecteert alle nodige gegevens op INSZ (data\_IDW) en transformeert alle absolute datums naar aantal kalenderdagen ten opzichte van de indexdatum van elke patiënt. Hierna wordt

elke INSZ vervangen door de corresponderende RN3 en worden de gegevens overgebracht naar de TTP VI (KSZ) (**stap 7**).

De FOD VVVL/MZG selecteert alle nodige gegevens op Hnew (data\_MZG). De opgevraagde data is afkomstig van patiënten met een ziekenhuisopname ten gevolge van een hart- en vaataandoening als hoofddiagnose tussen 01/01/2017 en 31/12/2018 (= indexdatum):

Verblijven & gegevens vanaf 1 jaar voor de indexdatum tot 2 jaar na de indexdatum:  
indexdatum-365dagen\* tot indexdatum+ 730 dagen\*

FOD VVVL/MZG transformeert vervolgens alle absolute datums naar aantal kalenderdagen ten opzichte van de indexdatum van elke patiënt.

Hierna wordt elke Hnew vervangen door de corresponderende RN1 en worden de gegevens overgebracht naar de TTP VI (KSZ) (**stap 8**).

Alle gegevens worden op Cproject geplaatst en overgebracht naar het IMA Project DWH (**stap 9**).

Een Small Cells Risk Analysis (SCRA) wordt uitgevoerd vóór de terbeschikkingstelling van de gegevens indien nodig geacht door het Informatieveilighheidscomité (IVC) (**stap 10**).

De gegevens van IDEWE, KSZ, FOD VVVL/MZG en IMA worden in het IMA Project DWH ter beschikking gesteld aan de onderzoekers (**stap 11**).

\* 366 dagen indien schrikkeljaar.

## **Bijlage 2: Overzicht van gevraagde data per gegevensbron**

- A. Toelichting en motivering
  - i. Algemeen

In kader van dit onderzoek naar de predictoren van werkhervatting na een Hart- en vaatziekte (HVZ) en de kost van HVZ, is het noodzakelijk om het ‘baseline’ zorggebruik, gebaseerd op zorggebruik respectievelijk 1 jaar voor de ziekenhuisopname, en het zorgtraject van de patiënten over een periode van twee jaar (2017-2019, 2018- 2020) na de ziekenhuisopname in kaart te brengen. Samengevat worden er gegevens opgevraagd van de volgende patiënten;

- Patiënten met een ziekenhuisopname ten gevolge van een hart- en vaataandoening als hoofddiagnose tussen 1/01/2017 en 31/12/2018 (= indexdatum\*\*):
  - Verblijven & gegevens vanaf 1 jaar voor de indexdatum tot 2 jaar na de indexdatum: indexdatum-365 dagen\* tot indexdatum+730 dagen\*

\* 366 dagen indien schrikkeljaar.

\*\* De indexdatum wordt bepaald aan de hand van de eerste ziekenhuisopname ten gevolge van HVZ als hoofddiagnose tussen 1/01/2017 en 31/12/2018.

Finaal wordt er een databestand per domein (DIAGNOSE, PROCEDURE, STAYHOSP, STAYXTRA, STAYSPEC, STAYINX) geleverd door FOD VVVL/MZG, waarbij elke rij een nieuwe verblijf (STAYNUM) voorstelt en de kolommen de aan het verblijf gekoppelde gegevens bevatten. Hierbij wordt een beperkt aantal variabelen opgevraagd voor elk nieuw verblijf (=STAYNUM) over de periode van 1 jaar voor de indexdatum tot 2 jaar na de indexdatum (zie tabel E.1.).

Alleen voor de index-opname (= HVZ als geverifieerde opnamediagnose van het eerste deelverblijf), worden daarnaast ook bijkomende variabelen opgevraagd in verband met deelverblijven, facturatiegegevens, procedures, etc. (zie tabel E.2.)

De datasets bevatten de gegevens van patiënten met een ziekenhuisopname ten gevolge van een hart- en vaataandoening als hoofddiagnose tussen 1/01/2017 en 31/12/2018. Deze exacte data worden uitsluitend gebruikt voor de selectie van de patiënten en bepaling van de indexdatum. De opgevraagde variabelen, gebaseerd op exacte data, worden steeds omgezet naar “aantal kalenderdagen ten opzichte van indexdatum” in het geleverde databestand.

- ii. Met betrekking tot aanvraag van socio-economische variabelen & werk-gerelateerde (psychosociale) variabelen.

Deze aanvraag kadert in een onderzoek naar de maatschappelijke kost van HVZ en tewerkstelling post-event (HVZ). Het inschatten van zowel directe als indirecte kosten zijn inherent bij dergelijke economische evaluaties vanuit het maatschappelijk perspectief. De sociaal-economische positie (bijvoorbeeld: arbeidsmarktpositie) laten toe om de indirecte kosten van zorg in te schatten. Daarnaast heeft de sociaal-economische positie -en de

evolutie hiervan in de tijd- een belangrijk effect op de verschillende sociale (financiële) beschermingsmaatregelen in België. De uitgebreidheid van socio-economische variabelen binnen dit onderzoek is noodzakelijk om die sociaal-economische posities en sociale beschermingsmaatregelen correct te bepalen zodoende de indirecte kosten correct in te schatten en op te volgen in de tijd. Daarnaast trachten de onderzoekers in dit onderzoek predictieve factoren voor werkhervatting en maatschappelijke kosten in kaart te brengen. Socio-demografische en -economische variabelen kunnen mogelijk een positieve of negatieve invloed uitoefenen op het al dan niet terugkeren naar het werk na een HVZ. Ook meer specifieke gegevens over het type werk en andere werk-gerelateerde factoren (IDEWE), spelen volgens de literatuur een grote rol bij tewerkstellingspredicties.

B. Variabelen die noodzakelijk zijn voor de selectie- en/of koppelingsdoeleinden

<b>Data- leverancier</b>	<b>Variabele</b>	<b>Bewijs van proportionaliteit</b>
<b>IMA</b>	SS00010 / PP0010 – Identificatie rechthebbende	Het INSZ-nummer/uniek identificatienummer zal enkel gebruikt worden voor selectie- en koppelingsdoeleinden. Deze variabele wordt omgezet naar een gepseudonimiseerd identificatienummer dat ter beschikking gesteld wordt voor de onderzoekers.
<b>KSZ</b>	INSZ – Identificatienummer van de sociale zekerheid	
<b>TCT-FOD (MZG)</b>	HNew	
<b>IDEWE</b>	INSZ	
<b>FOD (MZG)</b>	A2_CODE_DIAG_VE RIF_ADM  (ICD-10-BE (numeriek))	Deze variabele is noodzakelijk als selectievariabele voor het samenstellen van de cohort. Op basis van de combinatie van variabelen “Leeftijd”, “ICD-10 code (numeriek)” kunnen alle (arbeidsactieve) patiënten met een ziekenhuisopname ten gevolge van hart- en vaatziekte (HVZ) (hoofddiagnose) tussen 01/01/2017 en 31/12/2018. Concreet betreft het volgende diagnoses op basis van ICD-10-BE codes: I20-I25, I30, I33-I36, I38, I40, I46.0, I48, I49.0, I50, I60-I64, I71.

	<p>Incidentiedatum (HVZ) = Indexdatum</p> <p>FOD: A2_YEAR_HOSP_IN A2_MONTH_HOSP_IN A2_DAY_HOSP_IN &amp; A2_CODE_DIAG_VE RIF_ADM</p>	<p>De indexdatum wordt bepaald aan de hand van de eerste ziekenhuisopname ten gevolge van HVZ als hoofdiagnose tussen 01/01/2017 en 31/12/2018. Deze variabele is noodzakelijk als selectievariabele voor het samenstellen van de cohort. In kader van het dataminimalisatie-principe wordt geopteerd voor een methode waarbij geen exacte datums ter beschikking gesteld worden van de onderzoekers in de aangeleverde dataset. De exacte incidentiedatum wordt door FOD aangeleverd aan alle dataleveranciers opdat zij op hun beurt alle datumvariabelen kunnen omzetten naar 'aantal kalenderdagen t.o.v. indexdatum'. De lijst met incidentiedatums en patient- ID's worden niet bewaard door de dataleveranciers, maar worden bij elke nieuwe datalevering opnieuw aangeleverd door FOD waarbij eHealth codeersleutels kan voorzien. In geval van meer dan één ziekenhuisopname t.g.v. HVZ tijdens de indexperiode (van 01/01/2017 en 31/12/2018) geldt de eerste opname met hoofddiagnose HVZ als geverifieerde opname-diagnose (incidentiedatum) als indexdatum.</p> <p>De aangeleverde lijst met incidentiedatums omvat slechts één record per patiënt. De indexdatum wordt gefixeerd bij de eerste data-extractie. Nadien zullen er geen nieuwe patiënten toegevoegd kunnen worden aan de initiële selectie, het verwijderen van patiënten n.a.v. correcties is wel mogelijk. De conversiesleutel blijft steeds dezelfde over het volledige verloop van het project. Deze sleutels dienen -evenals de aangeleverde data - beschikbaar te blijven tot 10 jaar na de aanlevering van de dataset. Dit is noodzakelijk wanneer patiënten verwijderd zouden worden uit de dataset of in kader van bijkomende onderzoeksvragen.</p>
--	---	---

Tabel 1: Variabelen die noodzakelijk zijn voor koppelings- en/of selectiedoeleinden

### C. Selectie van variabelen vanuit de databank van IDEWE

<b>Dataleverancier</b>	<b>: IDEWE</b>
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie)	: 19

Variabele	Info/ Specificiëring	Bewijs van proportionaliteit
Onderworpen Werknemer	Werknemers die op medisch onderzoek komen omwille van beroepsrisico	Deze variabele stelt de onderzoekers in staat om de HVZ patiënten met beroepsrisico te identificeren, en om te controleren voor systematische missing data in de dataset.
Diagnose (ICD-9-CM)	Code ICD-9CM	Deze variabele vermeldt de ICD- 9 code voor de hoofd- en nevendiaagnoses tijdens de periode van ziekteverzuim. Deze variabele is noodzakelijk om het zorggebruik te linken aan ziekte. Zelfrapportage!
	Codelevel_twee	
	Codelevel_drie	



<b>Comorbiditeiten</b>	Comorbiditeit	Deze zelf gerapporteerde variabele maakt het mogelijk om secundaire, comorbide aandoeningen en klachten te identificeren. Deze variabele is een mogelijke voorspellende factor voor de kost van, en werkhervatting na HVZ.
<b>Lichamelijke klachten</b>	BMI	Deze variabelen geven een indicatie van de lichamelijke belasting, klachten en gezondheidsrisico's (tgv werk) waarmee een persoon geconfronteerd wordt. Tevens zijn zij markers voor (chronische) stress en cardiovasculair risico, wat mogelijk voorspellende variabelen zijn voor werkhervatting na HVZ
	Bloeddruk (systolisch & diastolisch)	
	Locomotorische klachten	
	Gezondheids- klachten tgv werk	
	Pijn op de borst	
	Evenwichts- problemen	
	Fysieke/ lichamelijke uitputting	
<b>Psychische klachten</b>	Onrust tgv werk	Deze variabelen geeft een indicatie van de psychische belasting (tgv werk) waarmee een persoon geconfronteerd wordt. Tevens is dit een markers voor (chronische) stress en cardiovasculair risico, wat mogelijk een voorspellende variabele is voor werkhervatting na HVZ
	Emotionele/ geestelijke uitputting	
<b>Job tevredenheid</b>		Deze kwalitatieve variabele geeft een indicatie van de mate waarin een persoon tevreden is met zijn job en is een mogelijke predictor voor werkhervatting.
<b>Classificatie werk (ISCO 08)</b>		Deze variabelen geven meer informatie weer betreffende de werksituatie en het type werk van een persoon. Deze kunnen noodzakelijk en verklarend zijn in de analysis betreffende voorspellende factoren van werkhervatting na een HVZ
Statuut		
<b>Zwaar beroep / Beroepsrisico</b>	Risicocodes: Ploegenarbeid, Nachtarbeid, Fysiek belastende job, Mentaal belastende job,...	Deze variabelen geven meer informatie weer betreffende de werksituatie en het beroepsrisico waar een persoon aan onderworpen wordt. Deze kunnen noodzakelijk en verklarend zijn in de analysis betreffende voorspellende factoren van werkhervatting na een HVZ.
Bedrijfsgrootte		Deze variabelen geven meer informatie weer betreffende de bedrijfsgrootte van een werkgever van een persoon. Deze kunnen noodzakelijk en verklarend zijn in de analysis betreffende voorspellende factoren vanwerkhervatting na een HVZ

Tabel 2: Selectie van variabelen vanuit de databank van IDEWE

Mogelijk onvolledig

Alleen voor Onderworpen Werknemers (OW) à Medisch onderzoek omwille van beroepsrisico

Vanaf 2020 bij OW (Persoonlijk medische vragenlijst (PMV)-AMH)

#### D. Selectie van variabelen vanuit de databank van het Intermutualistisch Agentschap

<b>Dataleverancier</b>	<b>: IMA</b>
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie) : 19 (Gezondheidszorgen), 10 (Farmanet), 40 (Populatiegegevens)	

Bestand(en)	Variabele	Bewijs van proportionaliteit
Gezondheidszorgen en Farmanet	SS00015 – Begindatum verstrekking / Datum van aflevering	Deze variabele laat toe om na te gaan wanneer een bepaalde verstrekking (voor de eerste keer) gefactureerd werd. Het betreft verpleegdagen, (toediening van) farmaceutische producten en prestaties. Deze datum wordt omgezet naar een gecodeerde datum ('aantal dagen t.o.v. indexdatum'). In combinatie met variabelen SS00020, SS00050 en SS00055 kunnen de directe zorgkosten op verschillende tijdsintervallen bepaald worden.
Gezondheidszorgen en Farmanet	SS00020 – Nomenclatuur / Code categorie geneesmiddel	De nomenclatuurcode of de categorie van het geneesmiddel (zoals voorgeschreven door het RIZIV) laten toe om na te gaan voor welke verstrekkingen de patiënt een vergoeding ontvangen heeft.
Gezondheidszorgen en Farmanet	SS00050 – Aantal gevallen / Hoeveelheid	Deze variabele geeft het aantal keer weer dat een verstrekking verleend is, het aantal leveringen/afgeleverde eenheden of hoeveelheden of het aantal keer dat een eenheidstarief aangerekend werd op een bepaalde dag/periode. Deze variabele laat toe om de totale kost voor een specifieke verstrekking te berekenen.
Gezondheidszorgen	SS00055 – Aantal dagen	Deze variabele geeft -vergelijkbaar met SS00050- het aantal gefactureerde dagen of forfaits weer in een bepaalde periode. Het IMA adviseert om beide variabelen op te nemen omdat de informatie niet altijd op correcte wijze geregistreerd wordt.
Gezondheidszorgen en Farmanet	SS00060 – ZIV-bedrag	Deze variabele geeft het bedrag weer dat terugbetaald werd door de verzekeringsinstelling aan de patiënt. Deze variabele vormt, in combinatie met variabelen SS00160 en SS00165, de belangrijkste bron van informatie over de directe medische kosten van het zorggebruik.
Gezondheidszorgen	SS00065(A) – Zorgverstrekkers-identificatiecode	Deze variabelen laten toe om de zorgverstrekker die verantwoordelijk is voor het uitvoeren van de prestatie te identificeren. Dit vormt een belangrijke bron van informatie omtrent het voorkomen van ziekte/aandoeningen. Variabele SS00065(A) wordt in gecodeerde vorm aangeleverd.

Gezondheidszorgen	SS00065(B) – Zorgverstreker-bekwaming	
Gezondheidszorgen	SS00075 – Identificatie instelling	Deze variabele is noodzakelijk om te traceren in welke instelling de rechthebbende van de verleende prestaties opgenomen is (of verblijft). Deze variabele wordt voorzien in een unieke gepseudonimiseerde code per voorziening.
Gezondheidszorgen	SS00080 – Dienstcode	Deze variabele laat toe om in detail de kostenplaats/dienstcode voor de verleende zorg binnen een bepaalde zorgvoorziening te bepalen.
Gezondheidszorgen	SS00085 – Plaats van verstrekking	Deze variabele laat toe om -op gecodeerde wijze- de instelling te identificeren waar de prestatie werkelijk uitgevoerd werd.
Gezondheidszorgen	SS00105 A – Identificatienummer derde SS00105 B – Derde bekwaming	Deze variabele laat toe om -op gecodeerde wijze- de instelling (type & erkenningsnummer) of verstreker die de prestatie gefactureerd heeft te identificeren indien er een derdebetalersregeling is.
Gezondheidszorgen	SS00110 – Datum opname SS00115 – Datum ontslag	Deze variabelen geven aan wanneer een patiënt opgenomen/ontslagen is uit een zorgvoorziening. Beide variabelen worden geconverteerd naar gecodeerde variabele ('aantal kalenderdagen t.o.v. indexdatum').
Gezondheidszorgen	<a href="#">SS00120-Type factuur</a>	Deze variabele geeft het type factuur dat werd opgemaakt weer, afhankelijk van de aard van de behandeling en het type van verblijf. De variabele is noodzakelijk om kosten die voortvloeien uit een bepaalde behandeling te kunnen verklaren.
Gezondheidszorgen	<a href="#">SS00130-Betrekkelijke verstrekking</a>	Deze variabele staat toe om een verstrekking te identificeren die afhankelijk is van een andere prestatie (bv. revalidatie na ingreep). Deze variabele laat toe om een duidelijker beeld te krijgen van het verband tussen verstrekkingen. Deze variabele is vereist om de variabele SS00150 te kunnen interpreteren.
Gezondheidszorgen en Farmanet	SS00135 – Nummer product	Deze code staat voor de code van het gefactureerde farmaceutisch product. Hierdoor kan exact bepaald worden welke geneesmiddelen de patiënt gebruikt. Hieruit kan o.a. afgeleid worden aan welke aandoeningen de patiënt lijdt.
Gezondheidszorgen	<a href="#">SS00145-Nacht, weekend of feestdag</a>	Deze variabele laat toe om de aangerekende supplementen voor verstrekkingen tijdens nacht, weekend of op een feestdag te identificeren.

Gezondheidszorgen	SS00150 – Gefactureerde nomenclatuurcodes	De gefactureerde nomenclatuurcode geeft de (pseudo-) nomenclatuurcode weer voor de verstrekking die de zorgverstrekker aangerekend heeft aan de patiënt. Het geeft informatie over pseudocodes waardoor de specifieke aard van de verstrekking bepaald kan worden.
Gezondheidszorgen en Farmanet	SS00160 – Remgeld	Deze variabele geeft het reëel persoonlijk aandeel (wettelijk remgeld) weer dat de patiënt heeft betaald voor de verstrekking. In combinatie met variabelen SS00060 en SS00165 kan hierdoor de directe medische kost berekend worden.
Gezondheidszorgen	SS00165 – Supplement	Deze variabele geeft informatie over supplementen die aangerekend werden aan de patiënt bovenop het persoonlijk aandeel. Deze supplementen vormen, in combinatie met variabelen SS00060 en SS00160, de basis voor de berekening van de directe medische kosten.
Farmanet	SS00165 – Vermindering verzekeringstegemoetkoming/Bijdrage van de apothekers	Deze variabele geeft de vermindering op de verzekeringstegemoetkoming van geneesmiddelen weer. Naar analogie van variabele SS00165 in het bestand Gezondheidszorgen, is deze variabele noodzakelijk om de kost van het geneesmiddelengebruik te berekenen.
Gezondheidszorgen	<a href="#">SS00170-Code implantaat</a>	De variabele geeft de identificatiecode weer van het implantaat of medisch invasief hulpmiddel (enkel opgenomen bij bepaalde gefactureerde nomenclatuurcodes). Deze variabele stelt de onderzoekers in staat om te bepalen of er bij een patiënt een implantaat werd ingebracht (Vb. Stent, Pacemaker,...) en is een mogelijk predictieve factor voor RTW.
Gezondheidszorgen	SS00175 – Derdebetalersregeling	Deze variabele laat toe om te bepalen of een patiënt al dan niet geniet van een voorkeursregeling volgens het derdebetalerssysteem. Dit kan relevant zijn voor patiënten die recht hebben op een verhoogde tegemoetkoming (identificeerbaar in variabele SS00160).
Farmanet	SS00195 – ZIV-tegemoetkoming 2	Deze variabele geeft de som weer van de aangerekende honoraria en laat een gedetailleerde berekening van de kost van het geneesmiddelengebruik toe.
Gezondheidszorgen	<a href="#">PROCEDURE AH CAT-Ambulant of gehospitaliseerd</a>	Deze variabele laat toe om samen met de variabele SS00020 na te gaan indien een verstrekking werd uitgevoerd in een gehospitaliseerde setting of niet.
Populatiegegevens	<a href="#">PP0020-Geslacht</a>	Deze variabele stelt de onderzoekers in staat het geslacht van de patiënt te bepalen en een clusteranalyse uit te voeren op basis van geslacht. Daarnaast is deze variabele een mogelijke predictor voor tewerkstelling na HVZ.

Populatiegegevens	PP0040 Jaar & maand overlijden	Deze variabele geeft weer wanneer een persoon met HVZ overlijdt na ontslag uit het ziekenhuis. Dit kan een belangrijke impact hebben op de medische kosten en de kosten t.g.v. productiviteitsverlies. Deze variabele wordt geconverteerd naar gecodeerde variabele ('aantal kalenderdagen t.o.v. indexdatum').
Populatiegegevens	<a href="#">PP1002-Gerechtigde of persoon ten laste</a>	Op basis van deze variabele kan men de hoedanigheid van een persoon in het mutualistisch dossier achterhalen ('Gerechtigde' of 'persoon ten laste'). Deze variabele laat toe om tewerkstelling en het sociaal statuut gelinkt aan de primaire behandeling in detail te bestuderen. Daarnaast is deze variabele een mogelijke predictor voor tewerkstelling na HVZ.
Populatiegegevens	<a href="#">PP1003-Sociaal statuut van de gerechtigde</a>	Deze variabele is belangrijk om het productiviteitsverlies correct in te schatten. De maatschappelijke kostprijs t.g.v. productiviteitsverlies zal namelijk verschillen wanneer het bijvoorbeeld arbeiders, bedienden, studenten of gepensioneerden betreft.
Populatiegegevens	<a href="#">PP1004-Type werkloosheidsuitkering</a>	Deze variabele geeft het type werkloosheidsuitkering weer. Deze variabele is van belang in het berekenen van de incrementele kost en laat toe exact te bepalen welke sociale maatregelen een impact hierop hebben.
Populatiegegevens	<a href="#">PP1008-Aard inkomen</a>	Deze variabele geeft een omschrijving van de aard van het inkomen van niet-actieve personen ingeschreven in het rijksregister. Deze variabele is noodzakelijk om de grootte van het productiviteitsverlies na traumatisch hersenletsel zo correct mogelijk in te schatten en is een mogelijke predictor voor werkhervatting.
Populatiegegevens	PP1009 - Oorsprong erkenning als mindervalide	Deze variabele is noodzakelijk om de eventuele dubbeltellingen rond toekenningen van ondersteuningsmaatregelen te identificeren en te vermijden
Populatiegegevens	PP1010 – Voorwaarde recht verhoogde tegemoetkoming	Deze variabele geeft weer of de patiënt recht heeft op een verhoogde tegemoetkoming (voorheen OMNIO-statuut). Deze variabele staat toe om het sociaal statuut gelinkt aan een behandeling in detail te bestuderen.
Populatiegegevens	PP2001 – Forfait B geneeskundige zorgen	Deze variabelen laten toe om te bepalen of de patiënt recht heeft op de omschreven tegemoetkoming/verstrekkingen. Deze variabele toe om het sociaal statuut gelinkt aan de primaire behandeling in detail te bestuderen. In kader van de berekening van de incrementele kost is het van belang om exact te bepalen welke sociale maatregelen een impact hebben op de vermindering in de directe en indirecte kosten van zorg <b>(tot 1 jaar pre- en 2 jaar post-indexdatum)</b> .
	PP2002 – Forfait C geneeskundige zorgen	
	PP2003 – Kinesitherapie E of fysiotherapie	

	PP2004 – Verhoogde kinderbijlagen	
	PP2005 – Integratietegemoetkoming voor gehandicapten	
	PP2006 – Tegemoetkoming hulp aan bejaarden	
	PP2007 – Tegemoetkoming voor hulp van derde (Oude regeling 1)	
	PP2008 – Tegemoetkoming voor hulp van derde (Oude regeling 2)	
	PP2009 –Tegemoetkoming voor hulp aan derden	
	PP2010 – Criterium hospitalisatie (120 dagen)	
	PP2011 – Criterium hospitalisaties (6 hospitalisaties)	
Populatiedata	PP3001 – Recht MAF – gezin	Deze variabelen laat toe te bepalen welk type MAF toegekend kan worden aan de patiënt en/of het gezin. Deze variabelen dienen samen geïnterpreteerd te worden.
	PP3002-Categorie MAF_gezin	
Populatiedata	PP3004 -MAF- Terugbetaling en plafond - gezin	Deze variabelen omschrijft de plafonnering die de basis vormt voor de hoogte van de terugbetaling.
Populatiedata	PP3005 – Terugbetaling en plafond - individu of gedeelte van het gezin	Deze variabele omschrijft het recht, en de plafonnering op basis waarvan de hoogte van de terugbetaling bepaald wordt.
Populatiedata	<a href="#">PP3010-Recht op gewaarborgd inkomen voor bejaarden, inkomensgarantie voor</a>	Deze variabelen laten toe om na te gaan indien een patiënt recht heeft op financiële steunmiddelen. In kader van de berekening van de incrementele kost is het van belang om exact te bepalen welke sociale maatregelen een impact hebben op de vermindering in de

	<a href="#">ouderen of leefloen</a>	directe en indirecte kosten van zorg.
	PP3011 – Recht op toelage voor gehandicapten	
	PP3013-OCMW-steun	
	PP3015-Statuut chronische aandoening – financieel criterium	
Populatiedata	PP4001 – Aantal dagen werkloosheid	Deze variabelen laten toe een inschatting te maken van het productiviteitsverlies. Deze variabele is noodzakelijk voor de berekening van de indirecte kosten. Daarnaast kunnen deze variabelen een mogelijke voorspeller zijn voor de tewerkstellingsstatus op 1 & 2 jaar post-event.
	PP4002 – Aantal dagen arbeidsongeschiktheid	
	PP4003 - Aantal dagen invaliditeit	
Populatiedata	<a href="#">PP4004-Graad van zelfredzaamheid</a>	Het is interessant om deze variabele mee te nemen in de analyse van werkhervatting na HVZ. Op deze manier kan worden nagegaan of zelfredzaamheid een predictieve waarde heeft m.b.t. werkhervatting
Populatiedata	<a href="#">PP4008-Eenoudergezin</a>	Deze variabelen laten toe om te bepalen of de patiënt recht heeft op de omschreven tegemoetkomingen/verstrekkingen. Deze variabele toe om het sociaal statuut gelinkt aan de primaire behandeling in detail te bestuderen en is een mogelijke predictor voor werkhervatting.
Populatiedata	PP0015-Geboortedatum (à Leeftijd)	Deze variabele wordt omgezet naar leeftijd (referentiejaar (indexjaar) – geboortjaar) en stelt de onderzoekers in staat om de leeftijd van de patiënt te bepalen op moment van ziekenhuisopname. Ze is noodzakelijk om de ‘arbeidsactieve’ bevolking te identificeren (18-65j) en voor het samenstellen van de cohort waarbij de onderzoekers tewerkstelling trachten te voorspellen.
Populatiedata (Flag)	FAM_SIZE-Gezinsgrootte	Deze variabele geeft het aantal leden in het MAF-gezin weer en is een mogelijke predictor voor werkhervatting.

Populatiedata	NIS code (Provincie & Arrondissement)	Aan de hand van de NIS-code kan men de woonplaats van de patiënt achterhalen. Op basis van de woonplaats kan bepaald worden aan welke regionale wetgeving een persoon onderhevig is. De variabele kan beperkt worden tot op niveau van de provincies indien arrondissementen een probleem zou vormen bij SCRA.
Populatiedata (Flag)	RESTHOME_DAYS	Deze variabele geeft aan hoeveel dagen de rechthebbende in de loop van het referentiejaar een forfait ROB / RVT / DVC heeft ontvangen. Ze vormt een belangrijke bepalende factor voor de kosten op basis van zorggebruik.
Populatiedata (Flag)	CHRONICAL_YN	Deze variabele laat toe om te bepalen of de patiënt in het afgelopen jaar recht had op een attest/statuut Chronische aandoening en/of een toelage voor gehandicapten. Deze variabele kan verklarend zijn voor verschillen in de kosten van zorg (als gevolg van de toekenning van verhoogde tegemoetkoming).
Populatiedata (Flag)	MAJOR_COVERAGE_YN	Deze flag laat toe om te bepalen of een patiënt al dan niet geniet van het voorkeurtarief verplichte verzekering of niet. Dit kan relevant zijn voor patiënten die recht hebben op een verhoogde tegemoetkoming
Populatiedata (Flag)	MAJOR_BENEFIT_YN	Deze flag geeft aan of de rechthebbende een w.i.g.w. (Weduwen, Invaliden, Gepensioneerden en Wezen) is (voorkeursregeling), of niet. Deze variabele is tevens een mogelijke predictor voor werkhervatting.
Populatiedata (Flag)	MAJOR_RISK_CAT	Deze flag geeft het hoofdstatuut van de verzekerde weer en geeft aan of de persoon (loontrekkende of zelfstandige) recht heeft op grote risico's.
Farmanet en Gezondheidszorgen (flag)	ATC_ANAT_L + ATC_THER_L + ATC_PHAR_L + ATC_CHEM_L n+ ATC_PROD_L	Deze flag is een afgeleide van variabele SS00135 (nummer product) volgende de referentietabellen van de ATC codes.
Farmanet en Gezondheidszorgen (flag)	PROCEDURE_AH_CAT PROCEDURE_GROUP PROCEDURE_DETAIL PROCEDURE_CAT	Deze afgeleide variabelen van ss00020 maken het gebruik van indelingen van procedures mogelijk zoals vastgelegd door het RIZIV
Gezondheidszorgen (flag)	PRACTITIONER_CAT	Deze variabele geeft het type van de verstrekker van de prestatie weer en maakt het mogelijk om te categoriseren/ clusteren per type zorgverlening die geraadpleegd werd.

Tabel 3: Selectie van variabelen vanuit de databank van het Inter mutualistisch Agentschap



E. Selectie van variabelen vanuit de databank van FOD VVVL (MZG databank)

Dataleverancier	: FOD MZG
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie)	: 36 opgevraagde variabelen (28 geleverde variabelen waarvan 4 gehercodeerde exacte datums) (18 variabelen opgevraagd voor alle verblijven gedurende baseline- & post-periode (=E.1.), 36 variabelen opgevraagd voor indexverblijf (= E.2.))

E.1. Selectie van variabelen voor ziekenhuisopnames in de baseline-periode (1 jaar pré- indexopname) & Post-periode (tot 2 jaar post-indexopname)

Bestand(en)	Variabele	Bewijs van proportionaliteit
STAYHOSP, STAYINDX, DIAGNOSE, PROCEDUR, PROCRIZI en TESTRESU	STAYNUM	Een nieuw verblijfsnummer wordt toegekend bij iedere ziekenhuisopname waardoor differentiatie tussen verleende zorg tijdens verschillende ziekenhuisopnames mogelijk is. De variabele vormt een belangrijke factor voor de interpretatie en analyse van zorgtrajecten en van de kostenbepaling op basis van zorggebruik.
DIAGNOSE, PROCEDUR, PROCRIZI en TESTRESU	CODE_AGR	Deze variabele geeft het erkenningsnummer van het ziekenhuis weer waardoor het mogelijk wordt om te bepalen waar (ziekenhuis) de zorg verleend werd . Ziekenhuiskenmerken, zijn geassocieerd met maten van ‘kwaliteit van zorg’, die op hun beurt geassocieerd zijn met werkhervatting [14, 22-24]. Deze variabele kan in combinatie met het aantal patienten geïnterpreteerd worden om ‘het volume’ van een ziekenhuis in te schatten, wat een mogelijke voorspeller is voor werkhervatting. Deze variabele wordt voorzien in een unieke gepseudonimiseerde code per voorziening.
STAYHOSP	A2_YEAR_HOSP_IN	Deze variabelen maken het mogelijk om de exacte opnameduur voor de gespecificeerde hospitalisatie (STAYNUM) te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: ‘aantal kalenderdagen t.o.v. indexdatum’. Deze variabele in combinatie met A2_HOSP_OUT is noodzakelijk om de duur van de hospitalisatie ten gevolge van een HVZ te bepalen, wat een voorspellende factor is voor tewerkstelling na HVZ [4].
	A2_MONTH_HOSP_IN	
	A2_DAY_HOSP_IN	
STAYHOSP	A2_YEAR_HOSP_OUT	Deze variabelen maken het mogelijk om de exacte opnameduur voor de gespecificeerde hospitalisatie (STAYNUM) te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: ‘aantal kalenderdagen t.o.v. indexdatum’. Deze variabele in combinatie met A2_HOSP_IN is noodzakelijk om de
	A2_MONTH_HOSP_OUT	

	A2_DAY_HOSP_OUT	duur van de hospitalisatie ten gevolge van een HVZ te bepalen, wat mogelijk een voorspellende factor kan zijn voor tewerkstelling na HVZ [4].
STAYHOSP	A2_HOSP_TYPE_FAC	Deze variabele geeft het verblijfstype weer op basis van facturatiegegevens met o.a. een onderscheid tussen “klassieke hospitalisatie” en “niet-chirurgische daghospitalisatie”. Dit is van belang om de zorgen verleend aan patiënten uit het cohort te bepalen en deze facturatiegegevens vormt een belangrijke factor voor de kostenbepaling op basis van zorggebruik.
STAYHOSP	A2_CODE_PLACE_BEFORE_ADM	Deze variabele geeft de plaats weer waar de patiënt zich bevond voor de ziekenhuisopname. De variabele is een mogelijke predictor voor tewerkstelling na een HVZ en speelt een verklarende rol bij predicties van tewerkstelling na een HVZ. ((Voorgaande) opname in internaat, gevangenis, verzorgingstehuis kan tewerkstelling beïnvloeden)
STAYHOSP	A2_CODE_ADM	Deze variabele geeft het type opname weer (o.a. “Geplande hospitalisatie” en “Dringende opname ...”). Deze codes kunnen van belang zijn om de urgentie van de behandelingen (en daaraan gerelateerde kosten) in te schatten, wat een mogelijke predictor is voor werkhervatting [25].
STAYHOSP	A2_CODE_DESTINATE	Deze variabele geeft weer waarnaar de patiënt verwezen wordt na ontslag. Onder de codes die toegekend worden o.a. “Overleden”. Deze code kan van belang zijn om te bepalen of de patiënt overleden is tijdens het verblijf in het ziekenhuis. Deze variabele speelt een belangrijke verklarende rol bij predicties van tewerkstelling post HVZ. (Indien “overleden”, geen RTW mogelijk).
STAYHOSP	A2_CODE_DISCHARGE	Deze variabele geeft of de patiënt ontslagen is uit het ziekenhuis (evt. op verwijzing). Onder de codes die toegekend worden o.a. “Overleden”. Deze code kan van belang zijn om te bepalen of de patiënt overleden is tijdens het verblijf in het ziekenhuis. Deze variabele speelt een belangrijke verklarende rol bij predicties van tewerkstelling post HVZ. (Indien “overleden”, geen RTW mogelijk)
STAYXTRA	Drg_36	De variabele diagnose gerelateerde groep (DRG) geeft de homogene pathologie groep weer van patiënten die naar verwachting vergelijkbare behoeften hebben, op basis van hun diagnoses, comorbiditeiten, behandelingen tot nu toe, en profielen (leeftijd,...). Deze variabele, in combinatie met Soi_36 en Rom_36 is een verklarende en bepalende factor in analyses naar zorggebruik, kosten en werkhervatting

	Soi_36	Deze variabele maakt het mogelijk om de ernst van de aandoening (severity of illness) die gepaard gaat met de Drg_36 juist in te schatten. Het kan een verklarende en belangrijke variabele zijn in analyses naar zorggebruik, kosten en werkhervatting
	Rom_36	Deze variabele maakt het mogelijk om het mortaliteitsrisico (risk of mortality) dat gepaard gaat met de Drg_36 juist in te schatten. Het kan een verklarende en belangrijke variabele zijn in analyses naar zorggebruik, kosten en werkhervatting
DIAGNOSE en PROCEDUR	TYPE_DIAGNOSE	Deze variabele maakt het mogelijk om de hoofd- en nevendiaagnoses van het eerste deelverblijf van ziekenhuisopname te differentiëren. De variabele is dus noodzakelijk voor de interpretatie van variabele "CODE_DIAGNOSE".
DIAGNOSE en STAYHOSP	CODE_DIAGNOSE A2_CODE_DIAG_VERIF- ADM	Deze variabele vermeldt de ICD code voor de hoofd- en nevendiaagnoses van de ziekenhuisopname. Deze variabele in combinatie met "TYPE_DIAGNOSE" is noodzakelijk om specifieke (comorbide) aandoeningen van patiënten binnen het cohort in kaart te brengen. De variabele maakt het tevens mogelijk om bepaalde subpopulaties te identificeren o.b.v. ICD-10 diagnoses. Daarnaast kan de combinatie van bepaalde hoofd- en nevendiaagnoses een indicatie zijn voor de complexiteit van de aandoening, wat mogelijk een voorspeller is voor werkhervatting [4,14]

**E.2. Selectie van variabelen bij de indexopname (Hart & vaatziekte als geverifieerde opname (Hoofd-) diagnose van eerste deelverblijf tussen 01/01/2017 en 31/12/2018)**

Bestand(en)	Variabele	Bewijs van proportionaliteit
STAYHOSP, STAYINDEX, DIAGNOSE, PROCEDUR, PROCRIZI en TESTRESU	STAYNUM	Een nieuw verblijfsnummer wordt toegekend bij iedere ziekenhuisopname, waardoor differentiatie tussen verleende zorg tijdens verschillende ziekenhuisopnames mogelijk is. De variabele vormt een belangrijke factor voor de interpretatie en analyse van zorgtrajecten en van de kostenbepaling op basis van zorggebruik.
PATHOSPI, STAYXTRA	Leeftijd A1_YEAR BIRTH	Deze variabele stelt de onderzoekers in staat om de leeftijd van de patiënt te berekenen op het moment van de ziekenhuisopname. Deze is noodzakelijk om de 'arbeidsactieve' bevolking te identificeren (18-65j) en voor het samenstellen van de cohort waarbij de onderzoekers tewerkstelling trachten te voorspellen. Daarnaast is leeftijd een voorspellende factor voor tewerkstelling na HVZ [4].

DIAGNOSE, PROCEDUR, PROCRIZI en TESTRESU	CODE_AGR	Deze variabele geeft het erkenningsnummer van het ziekenhuis weer waardoor het mogelijk wordt om te bepalen waar (ziekenhuis) de zorg verleend werd. Ziekenhuiskenmerken, zijn geassocieerd met maten van ‘kwaliteit van zorg’, die op hun beurt geassocieerd zijn met werkhervatting [14, 22-24]. Deze variabele kan in combinatie met het aantal patiënten geïnterpreteerd worden om ‘het volume’ van een ziekenhuis in te schatten, wat een mogelijke voorspeller is voor werkhervatting. Deze variabele wordt voorzien in een unieke gepseudonimiseerde code per voorziening.
STAYHOSP	A2_YEAR_HOSP_IN	Deze variabelen maken het mogelijk om de exacte opnameduur voor de gespecificeerde hospitalisatie (STAYNUM) te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: ‘aantal kalenderdagen t.o.v. indexdatum’.
	A2_MONTH_HOSP_IN	
	A2_DAY_HOSP_IN	Deze variabele in combinatie met A2_HOSP_OUT is noodzakelijk om de duur van de hospitalisatie ten gevolge van een HVZ te bepalen, wat een voorspellende factor is voor tewerkstelling na HVZ [4].
STAYHOSP	A2_YEAR_HOSP_OUT	Deze variabelen maken het mogelijk om de exacte opnameduur voor de gespecificeerde hospitalisatie (STAYNUM) te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: ‘aantal kalenderdagen t.o.v. indexdatum’. Deze variabele in combinatie met A2_HOSP_IN is noodzakelijk om de duur van de hospitalisatie ten gevolge van een HVZ te bepalen, wat mogelijk een voorspellende factor kan zijn voor tewerkstelling na HVZ [4].
	A2_MONTH_HOSP_OUT	
	A2_DAY_HOSP_OUT	
STAYHOSP	A2_HOSP_TYPE_FAC	Deze variabele geeft het verblijfstype weer op basis van facturatiegegevens met o.a. een onderscheid tussen “klassieke hospitalisatie” en “niet-chirurgische daghospitalisatie”. Dit is van belang om de zorgen verleend aan patiënten uit het cohort te bepalen en deze facturatiegegevens vormt een belangrijke factor voor de kostenbepaling op basis van zorggebruik.
STAYHOSP	A2_CODE_PLACE_BEFORE_ADM	Deze variabele geeft de plaats weer waar de patiënt zich bevond voor de ziekenhuisopname. De variabele is een mogelijke predictor voor tewerkstelling na een HVZ en speelt een verklarende rol bij predicties van tewerkstelling na een HVZ. ((Voorgaande) opname in internaat, gevangenis, verzorgingstehuis kan tewerkstelling beïnvloeden)
STAYHOSP	A2_CODE_ADM	Deze variabele geeft het type opname weer (o.a. “Geplande hospitalisatie” en “Dringende opname ...”). Deze codes kunnen van belang zijn om de urgentie van de behandelingen (en daaraan gerelateerde kosten) in te schatten, wat een mogelijke predictor is voor werkhervatting [25].

STAYHOSP	A2_CODE_DESTINATE	Deze variabele geeft weer waarnaar de patiënt verwezen wordt na ontslag. Onder de codes die toegekend worden o.a. “Overleden”. Deze code kan van belang zijn om te bepalen of de patiënt overleden is tijdens het verblijf in het ziekenhuis. Deze variabele speelt een belangrijke verklarende rol bij predicties van tewerkstelling post HVZ. (Indien “overleden”, geen RTW mogelijk)
STAYHOSP	A2_CODE_DISCHARGE	Deze variabele geeft of de patiënt ontslagen is uit het ziekenhuis (evt. op verwijzing). Onder de codes die toegekend worden o.a. “Overleden”. Deze code kan van belang zijn om te bepalen of de patiënt overleden is tijdens het verblijf in het ziekenhuis. Deze variabele speelt een belangrijke verklarende rol bij predicties van tewerkstelling post HVZ. (Indien “overleden”, geen RTW mogelijk)
STAYHOSP	A2_CODE_SEX	Als een standaard demografisch aspect is het belangrijk een onderscheid te maken naar geslacht bij het beschrijven van ziektelast, zorggebruik, kosten en werkhervatting. Ook als mogelijke verklarende factor is deze variabele relevant. Deze variabelen kunnen voorspellend zijn voor ziektelast, zorggebruik, kosten en werkhervatting.
STAYHOSP	A2_CODE_STAT_INSURANCE	Deze code geeft de verzekeringsstatus van de patiënt weer tijdens het verblijf. Met deze variabele kan men nagaan of een patiënt (van Belgische nationaliteit of andere nationaliteit) aangesloten is bij een Belgische verzekering. Deze variabele is noodzakelijk om eventuele missing data uit de IMA- dataset te interpreteren. Daarnaast benadrukt voorgaand onderzoek [21], dat nationaliteit bepalend is voor de duur en kans op werkhervatting.
STAYHOSP	A2_CODE_INDIC_NAT	Via deze variabele kan nagegaan worden welke nationaliteit een patiënt met hersentrauma heeft. Een deel van de buitenlandse betrokkenen (diegenen zonder Belgisch rijksregisternummer) zullen niet teruggevonden worden in de databanken van het IMA. Deze variabele, in combinatie met A2_CODE_STAT_INSURANCE kan dus verklarend zijn voor een aantal missing links in de koppeling. Daarnaast benadrukt voorgaand onderzoek [21], dat nationaliteit bepalend is voor de duur en kans op werkhervatting.
STAYSPEC	CODE_SPEC	Deze variabelen geven de verschillende specialisten (en hun volgnummer) weer, die verantwoordelijk zijn voor de hoofd- en nevendiaagnoses tijdens een deelverblijf van de gespecificeerde hospitalisatie. De variabele vormt een belangrijke factor voor de kostenbepaling op basis van zorggebruik. Het aantal specialisten die een patiënt ziet, is mogelijk gerelateerd aan complexiteit van de aandoening wat kan bepalend zijn voor de kans op RTW [4,14].
	ORDER_SPEC	

STAYXTRA	Drg_36	De variabele diagnose gerelateerde groep (DRG) geeft de homogene pathologie groep weer van patiënten die naar verwachting vergelijkbare behoeften hebben, op basis van hun diagnoses, comorbiditeiten, behandelingen tot nu toe, en profielen (leeftijd,...). Deze variabele, in combinatie met Soi_36 en Rom_36 is een verklarende en bepalende factor in analyses naar zorggebruik, kosten en werkhervatting
	Soi_36	Deze variabele maakt het mogelijk om de ernst van de aandoening (severity of illness) die gepaard gaat met de Drg_36 juist in te schatten. Het kan een verklarende en belangrijke variabele zijn in analyses naar zorggebruik, kosten en werkhervatting
	Rom_36	Deze variabele maakt het mogelijk om het mortaliteitsrisico (risk of mortality) dat gepaard gaat met de Drg_36 juist in te schatten. Het kan een verklarende en belangrijke variabele zijn in analyses naar zorggebruik, kosten en werkhervatting
STAYINDEX en PROCRIZI	ORDER_BEDINDEX	Deze variabele geeft in chronologische volgorde de bedindexen (dienstcodes / kostenplaatsen) weer waar de patiënt verbleven heeft tijdens de gespecificeerde hospitalisatie. Deze code in combinatie met “A4_YEAR/MONTH/DAY_BEDINDEX_IN” is van belang om de zorgtrajecten van de cohort in kaart te brengen en een onderscheid te maken tussen de verschillende bedindexen binnen een ziekenhuisopname. De variabele vormt daarnaast een belangrijke factor voor de kostenbepaling op basis van zorggebruik.
STAYINDEX	A4_YEAR_BEDINDEX_IN	Deze variabelen maken het mogelijk om de opnamedag op een bepaalde bedindex tijdens een gespecificeerde hospitalisatie te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: ‘aantal kalenderdagen t.o.v. indexdatum’. De variabele in combinatie met “ORDER_BEDINDEX” is noodzakelijk om tijdens de hospitalisatie de duur van het verblijf per bedindex te berekenen. Dit geeft een indicatie van het zorgtraject, de urgentie en complexiteit van de aandoening, wat een mogelijke voorspeller is van werkhervatting na HVZ [4,14].
	A4_MONTH_BEDINDEX_IN	
	A4_DAY_BEDINDEX_IN	
STAYINDEX	A4_NUMBER_DAY_FAC	Deze variabele geeft het aantal aaneensluitende dagen op een bepaalde bedindex weer en is van belang om de trajecten van de cohort in kaart te brengen. Zo kunnen er onderbrekingen tijdens de ziekenhuisopnames worden vastgesteld. Deze facturatiegegevens zijn noodzakelijk voor de kostenbepaling op basis van zorggebruik
STAYINDEX	A4_NUMBER_DAY_FAC_PREV	Deze variabele laat toe om te corrigeren voor gefactureerde dagen in het voorgaande registratiejaar. Deze variabele dient samen geïnterpreteerd te worden met variabele “A4_NUMBER_DAY_FAC”. En is een belangrijke factor voor de kostenbepaling op basis van zorggebruik

DIAGNOSE PROCEDUR STAYSPEC	ORDER_SPEC	Deze variabele geeft het volgnummer van de verschillende specialisten weer die verantwoordelijk zijn voor de hoofd- en nevendiaagnoses tijdens een gespecificeerde hospitalisaties.
DIAGNOSE en PROCEDUR	TYPE_DIAGNOSE	Deze variabele maakt het mogelijk om de hoofd- en nevendiaagnoses van het eerste deelverblijf van ziekenhuisopname te differentiëren. De variabele is dus noodzakelijk voor de interpretatie van variabele “CODE_DIAGNOSE”.
DIAGNOSE en STAYHOSP	CODE_DIAGNOSE A2_CODE_DIAG_VERIF_ ADM	Deze variabele vermeldt de ICD code voor de hoofd- en nevendiaagnoses van het eerste deelverblijf van ziekenhuisopname. Deze variabele is noodzakelijk voor de identificatie van relevante patiënten binnen het cohort en is nodig om het zorggebruik te linken aan ziekte, (zie selectievariabelen: bijlage 3 B).
PROCEDUR	M2_CODE_PROCEDURE	Deze variabele omschrijft de code van de uitgevoerde ingrepen tijdens de indexopname (A2_CODE_DIAG_VERIF_ADM = ICD-10-BE codes: I20-I25, I30, I33- I36, I38, I40, I46.0, I48, I49.0, I50, I60-I64, I71). Deze variabele is noodzakelijk om uitgevoerde ingrepen (verleende zorgen) en het behandelingsprogramma te bepalen, wat bepalend is voor de kans op werkhervatting [4,11]
PROCEDUR	M2_YEAR_PROCEDURE	Deze variabelen laten toe om de datum van de uitgevoerde ingrepen te bepalen. De combinatie van deze drie variabele wordt omgezet naar één gecodeerde variabele: ‘aantal kalenderdagen t.o.v. indexdatum’. De procedures stellen de onderzoekers in staat om de behandelingsprogramma’s voor patiënten in het cohort in kaart te brengen, wat bepalend kan zijn voor de kans op werkhervatting [4,11].
	M2_MONTH_PROCEDURE	
	M2_DAY_PROCEDURE	
PROCEDUR	M2_NUMBER_ PROC_1_DAY	Deze variabele geeft aan of de ingreep meerdere keren uitgevoerd werd op dezelfde dag. de variabele vormt een belangrijke factor voor de kostenbepaling op basis van zorggebruik en is noodzakelijk om kosten die voortvloeien uit een bepaalde ingreep te kunnen verklaren.

Tabel 4: Selectie van variabelen vanuit de databank van de Federale Overheidsdienst Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu (Minimale Ziekenhuis Gegevens)

F. Selectie van variabelen vanuit de databank van KSZ

<b>Dataleverancier</b>		<b>: KSZ</b>
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie)		: 63
<b>Bron</b>	<b>Variabele</b>	<b>Bewijs van proportionaliteit</b>
DWH_BCSS_NomenclatureVarD er	Nomenclatuur van socio-economische positie (Nomenc)	<p>Deze variabele laat toe om de socio-economische positie (tewerkstellingsvorm) te bepalen. Via deze variabele kan een inschatting gemaakt worden van het percentage en de snelheid van werkhervatting na ziekte.</p> <p>Dit is noodzakelijk om een inschatting te maken van de indirecte kosten van zorg (in functie van doelstelling 2). Deze variabele wordt geïnterpreteerd in combinatie met de variabelen van “bijkomende socio- economische positie” (zie hieronder).</p> <p>Om de her-identificatie van personen in de dataset te verhinderen, wordt het detail van de variabelen beginnend met 1 (werkend) beperkt tot dit eerste niveau.</p>
	Bijkomende positie van socio-economische positie ( <i>hercodering</i> )	<p>Om de her-identificatie van personen in de dataset te verhinderen, werden een aantal variabelen gehercodeerd waarbij het plausibel geacht wordt dat deze personen administratief werkend zijn, doch feitelijk hun beroepsactiviteit (tijdelijk) stopgezet hebben. Concreet betreft het volgende variabelen:</p> <ul style="list-style-type: none"> <li>- Werkend en gedeeltelijk in terbeschikkingstelling voorafgaand aan pensioen (Werkend_en_gedeeltelijk_TBS)</li> <li>- Werkend en gedeeltelijke loopbaanonderbreking/deeltijds tijds krediet (interr_carriere_temps_partie1)</li> </ul>
		<ul style="list-style-type: none"> <li>- Werkend en halftijds in een stelsel van werkloosheid met bedrijfstoeslag (prepension_mi_temps)</li> <li>- Werkende en volledig in terbeschikkingstelling voorafgaand aan pensioen (Werkend_en_volledig_TBS)</li> <li>- Werkend en volledige loopbaanonderbreking/volledig tijds krediet (interr_carriere_complete)</li> <li>- Werkend en volledig in stelsel van werkloosheid met bedrijfstoeslag (prepension_complete)</li> <li>- Werkend als pensioentrekkende (werkend_als_pensioentrekkende)</li> <li>- Invaliditeit en werkend (riziv_travailleur)</li> </ul>
	Activation_ONEm (Activering RVA)	De variabele geeft de Werkende arbeidsmarktpositie weer in combinatie met een activering van de RVA. Deze extra activering heeft mogelijk implicaties voor de tijd tot werkhervatting.



DWH_BCSS_Revenu	loon_BB_rsz	Deze variabelen, ingedeeld in klassen per 100 euro, zijn noodzakelijk om het bruto belastbaar inkomen te bepalen. Deze variabelen bevatten zowel inkomen uit arbeid als inkomen uit uitkeringen. De variabele in combinatie met het opleidingsniveau van de patiënt, geeft een indicatie van de sociaal economische status, wat een mogelijke voorspeller is voor werkhervatting na HVZ.	
	loon_BB_rszppo		
	uitkering_BB_DG_kg		
	uitkering_BB_FAMIFE D		
	uitkering_BB_fao		
	uitkering_BB_fbz	Meer bepaald vragen de onderzoekers loon- & uitkeringsklassen op van 1 jaar pré-index tot en met 2 jaar post- index.	
	uitkering_BB_fod_sz		
	uitkering_BB_nic		
	uitkering_BB_pod_mi		Voor een patiënt met indexdatum = 01/04/2017 betekent dit dat de onderzoekers de ‘bedragsklasse’ opvragen; op 31/12/2016 (= Loon 1 jaar pré-index, 2016), op 31/12/2017 (= Loon van indexjaar, 2017), op 31/12/2018 (= Loon 1 jaar post-index,2018) en op 31/12/2019 (= Loon 2 jaar post-index,2019)
	uitkering_BB_riziv		
	uitkering_BB_rva		
	uitkering_BB_rvp		
uitkering_BB_vsb			
DWH_STATBEL_Education	Onderwijsniveau (EDU)	Deze variabele geeft een momentopname van het hoogst behaalde opleidingsniveau van een individu tussen 2012 en 2017. Om echter na te gaan of de personen in het cohort een hoger onderwijsniveau behaalden in de jaren volgend op deze periode dienen onderstaande variabelen (Opleidingsniveau, cat_shu, cat_études, CO_DIPLOME, Studieniveau) eveneens opgevraagd te worden.	
DWH_VDAB_FOR EM_ACTIRIS_ ADG	Opleidingsniveau (Opleidingsniveau)	Deze variabele geeft het opleidingsniveau weer (Vlaamse gemeenschap), en kan verklarend zijn voor de snelheid van werkhervatting.	
DWH_ONEm_FIN_ DROIT_ART3 6	Studieniveau (Studieniveau)	Deze variabele beschrijft het studieniveau van individuen die werkloos zijn ten gevolge van het vervallen van het recht op inschakelingsuitkering. Deze variabele kan verklarend zijn voor de snelheid van werkhervatting.	
DWH_CREF_diplomes	Studiecategorie (niet-universitair onderwijs) (cat_shu)	Deze variabele geeft het onderwijsniveau weer (Franstalige gemeenschap), en kan verklarend zijn voor de snelheid van werkhervatting na een periode van ziekte.	
	Studiecategorie (cat_études)		

	Code van diploma (CO_DIPLOME)		
DWH_ONSS_StatbaseDMFA (tot 2016)	Prestatietype (T_Prest)	Deze variabele geeft het arbeidsregime (voltijds, deeltijds, speciaal, onbepaald, afwezigheid wegens ziekte) op de laatste dag van het kwartaal weer. Deze variabele laat toe om te bepalen of (en in welk regime) de patiënt werkzaam is.	
	Percentage deeltijds (Tauxpt)	Het percentage van een voltijdse job dat een deeltijdse werknemer presteert.	
	Voltijdsequivalent exclusief gelijkgestelde dagen (Eqtpsp)	Deze variabele geeft het tewerkstellingspercentage per arbeidsplaats weer gedurende het kwartaal.	
	Percentage Voltijdsequivalent exclusief gelijkgestelde dagen (Tauxsa)	Deze variabele geeft het tewerkstellingspercentage per arbeidsplaats weer gedurende het kwartaal in een percentage.	
	Voltijdsequivalent inclusief gelijkgestelde dagen (Tauxaa)	Deze variabele geeft het tewerkstellingspercentage per arbeidsplaats weer gedurende het kwartaal.	
	Codtra (Werknemerscode)	Deze variabelen geven informatie over tewerkstelling binnen de sociale economie. Dit is naast het normale economische circuit een mogelijke uitkomst op het vlak van werkhervatting.	
	DWH_ONSS_UniStatbaseDMFA (vanaf 2017)	Clatra (Werknemersklasse)	Verschillende studies hebben reeds uitgewezen dat het onderscheid tussen arbeiders en bedienden mogelijk een relevante predictor van werkhervatting na HVZ is.
		Clatr2 (Werknemersklasse speciaal)	Deze variabele geeft informatie omtrent een eventuele tewerkstelling binnen de sociale categorie.
		R_exclus Indicatorveld arbeidsprestatie	Deze code duidt aan of een record moet worden meegeteld of niet en is dus vereist voor een correcte interpretatie van de data.
		Secemp (sector)	Deze variabele geeft aan of een persoon tewerkgesteld is in de private dan wel de publieke sector. Er zal worden nagegaan of er mogelijks een verband is tussen deze variabele en de kans op arbeidsre-integratie
Belangrijkeheidscode (Codimp)		Deze variabelen geven meer informatie weer betreffende de werksituatie en het type werk van een persoon. Deze kunnen noodzakelijk en verklarend zijn in de analyse betreffende voorspellende factoren van werkhervatting na een HVZ.	
NaceEtab			
DWH_INASTI_RGT I	NACE-code (Nacec)		

DWH_CIN	Begindatum arbeidsongeschiktheid (Begindatum)	Deze variabele geeft de begin- en einddatum van arbeidsongeschiktheid weer. Deze variabelen worden geconverteerd naar twee gecodeerde variabelen ('aantal kalenderdagen t.o.v. indexdatum'). Deze variabele laat toe om het productiviteitsverlies te bepalen. De variabele is noodzakelijk voor de berekening van de indirecte kosten. De exacte berekening van indirecte kosten zijn inherent bij het bepalen van de maatschappelijke kost.
	Einddatum arbeidsongeschiktheid (Einddatum)	
	Aantal dagen arbeidsongeschiktheid (Aantal_dagen_ao)	Deze variabele wordt beschouwd als een controlevariabele op de variabelen "Begindatum" en "Einddatum". Deze variabele maakt het mogelijk om op eenvoudige wijze berekeningen te maken omtrent productiviteitsverlies en indirecte kosten.
	Samenstelling gezin (Samenst_gezin)	Deze variabele geeft de samenstelling van het gezin weer en kan verklarend zijn voor het toekennen van verhoogde tegemoetkomingen.
	Srt_dagen (soort dagen)	Met deze variabele kunnen redenen van arbeidson- geschiktheid uitgesloten worden die niet het gevolg kunnen zijn van HVZ (bv. Moederschapsrust). Zo worden verkeerde interpretaties voorkomen.
DWH_INAMI_CMI	Type ziekte (Type_ziekte)	Deze variabele is noodzakelijk om variabele "Diag_ziekte" te interpreteren. Deze variabele geeft het coderingssysteem (ICD-9 of ICD-10) weer waarin de ziekte gecodeerd werd.
	Medische code (Gr_ziekte)	Deze variabele laat toe om de oorzaak van invaliditeit (erkend door Geneeskundige Raad voor Invaliditeit) te bepalen. Dit laat toe om de invaliditeit te linken aan een behandeling.
	Diag ziekte (Diag_ziekte)	Deze variabele vermeldt de ICD code op basis waarvan de invaliditeit (variabele "Gr_ziekte") toegekend werd.
	Begin (B_ziekte)	Deze variabele geeft de begindatum van de primaire arbeidsongeschiktheid weer. Deze variabele wordt omgezet naar een geconverteerde variabele ('aantal kalenderdagen t.o.v. indexdatum'). Deze variabele wordt beschouwd als een controlevariabele op de variabele "Begindatum".
	Begindatum [invaliditeit] (B_inval)	Deze variabele geeft de begin- en einddatum van invaliditeit weer. Deze variabelen worden geconverteerd naar twee gecodeerde variabelen ('aantal kalenderdagen t.o.v. indexdatum'). Deze variabele laat toe om het productiviteitsverlies te bepalen. Deze variabele is noodzakelijk voor de berekening van indirecte kosten. De exacte berekening van indirecte kosten zijn inherent bij het bepalen van de maatschappelijke kost.
	Einddatum [invaliditeit] (E_inval)	

<a href="#">DWH_INAMI_CMI_RELAPSE:</a>	Indicatorveld invaliditeit R_exclus	Deze variabele is noodzakelijk om te corrigeren bij het samenvoegen van data uit verschillende kwartalen.
	d_relapse	Deze variabele geeft de datum van herval in invaliditeit weer. Deze variabele wordt geconverteerd naar een gecodeerde variabelen ('aantal dagen t.o.v. indexdatum'). Deze variabele laat toe om het productiviteitsverlies te bepalen.en is noodzakelijk voor de berekening van indirecte kosten
DWH_INAMI_Paiements	Bedrag uitkering (Bedrag)	Deze variabelen laten toe om het totale uitgekeerde bedrag te berekenen voor een vermelde periode op basis van de combinatie 'Vergoede_dagen' en 'Bedrag'. Deze variabelen zijn samen met de variabele 'IncapacityCategory' noodzakelijk om een exacte berekening uit te voeren voor de indirecte en maatschappelijke kosten.
	Uitkeringsdagen (Vergoede_dagen)	
	IncapacityCategory (code)	
	Duur werkloosheid (DUUR)	Deze variabele geeft de duur van en werkloosheid weer voor een individu en laten toe een inschatting te maken van het productiviteitsverlies. Deze variabele is noodzakelijk voor de berekening van de indirecte kosten.
DWH_RN_EtatCivil	Burgerlijke staat (Burgerlijke_staat)	Deze variabelen geven meer informatie betreffende de sociale staat van een individu in het cohort en kunnen voorspellend of verklarend zijn in de analyses voor maatschappelijke kosten/of werkhervatting.
DWH_RN_BCSS_BisTer	Type huishouden (Type_huishouden)	
	LIPRO (LIPRO-positie)	
DWH_AHOVOKS_LED	ISCED_niveau	Uit de literatuur omtrent de voorspelling van werkhervatting na HVZ blijkt dat het opleidingsniveau van de betrokkene een bepalende factor is voor re-integratie op de arbeidsmarkt. Daarom is het belangrijk deze informatie te kunnen opnemen in het predictiemodel.

Tabel 5: Selectie van variabelen vanuit de databank van de Kruispuntbank van de Sociale Zekerheid